UNIVERSIDADE ESTADUAL PAULISTA
"JÚLIO DE MESQUITA FILHO"
Campus de São José do Rio Preto

Raul Tintaya Marcavillaca

Condições de Otimalidade para Controle Ótimo via Formalismo de Dubovitskii-Milyutin

Dissertação de Mestrado
Pós-Graduação em Matemática

Instituto de Biociências, Letras e Ciências Exatas
Rua Cristóvão Colombo, 2265, 15054-000
São José do Rio Preto - SP - Brasil
Telefone: (17) 3221-2444 - Fax: (17) 3221-2445

Raul Tintaya Marcavillaca

Condições de Otimalidade para Controle Ótimo via Formalismo de Dubovitskii-Milyutin

Orientador:
Prof. Dr. Valeriano Antunes de Oliveira

Universidade Estadual Paulista "Júlio de Mesquita Filho"
Instituto de Biociências, Letras e Ciências Exatas
Campus de São José do Rio Preto

São José do Rio Preto

03 de Março de 2016

Tintaya M., Raul.

Condições de Otimalidade para Controle Ótimo via Formalismo de Dubovitskii-Milyutin / Raul Tintaya Marcavillaca. – São José do Rio Preto, 2016

109 f.

Orientador: Valeriano Antunes de Oliveira

Dissertação (mestrado) - Universidade Estadual Paulista “Júlio de Mesquita Filho”, Instituto de Biociências, Letras e Ciências Exatas

1. Matemática. 2. Teoria do controle. 3. Otimização matemática. 4. Princípios do máximo (matemática). I. Oliveira, Valeriano Antunes de. II. Universidade Estadual Paulista “Júlio de Mesquita Filho”, Instituto de Biociências, Letras e Ciências Exatas. III. Título.

CDU - 517.91.01

Ficha catalográfica elaborada pela Biblioteca do IBILCE

UNESP - Câmpus de São José do Rio Preto

Raul Tintaya Marcavillaca

Condições de Otimalidade para Controle Ótimo via Formalismo de Dubovitskii-Milyutin

Dissertação apresentada como parte dos requisitos para obtenção do título de Mestre em Matemática, junto ao Programa de Pós-Graduação em Matemática, do Instituto de Biociências, Letras e Ciências Exatas da Universidade Estadual Paulista "Júlio de Mesquita Filho", Campus de São José do Rio Preto.

BANCA EXAMINADORA

Prof. Dr. Valeriano Antunes de Oliveira
UNESP - São José do Rio Preto
Orientador

Prof. Dr. Geraldo Nunes Silva
UNESP - São José do Rio Preto

Prof. Dr. Iguer Luis Domini dos Santos
UNESP - Ilha Solteira

São José do Rio Preto, 03 de Março de 2016.

A meus pais Eusebio e Gerarda.

A meus queridos avós e irmãos.

Dedico.

AGRADECIMENTOS

Em primeiro lugar agradeço a Deus por ter me dado a vida e ter me permitido chegar até aqui.

A meus pais Eusebio e Gerarda por todo o carinho e apoio incondicional. Por ter acreditado sempre em mim e por todo os esforços que eles fizerem para eu estar aqui.

Agradeço a meus irmãos Eber, Jackely e Charmely Lisbeth por seus incentivos e por todos os momentos que passamos juntos.

Agradeço infinitamente ao professor Dr. Valeriano Antunes de Oliveira, meu orientador, pela confiança, dedicação, sugestões, correções, pela infinita paciência durante a elaboração deste trabalho.

A meus professores da UNSAAC, Professor Guido Alvarez, Patrício Choque e Alejandro Ttito pelo apoio e incentivo durante a graduação.

Agradeço também a meus amigos Gino, Daniel, Ismael, Paola, Daniela, Diana, Marco, Angelica, Rodiak e Fabíola pelos conselhos e amizade. Agradeço especialmente a Jacqueline pela compreensão, amizade e companhia, e pela ajuda e dicas com o Latex.

Aos professores do Departamento acadêmico de Matemática Pura e Aplicada e funcionários do IBILCE que, direta ou indiretamente, contribuíram para a elaboração deste trabalho.

Ao CNPq pelo auxílio financeiro.

A todos aqueles que de alguma maneira contribuíram para este trabalho, mas por algum motivo não foram citados.

Faça as coisas o mais simples que puder,
porém não as mais simples.
Albert Einstein

RESUMO

O objetivo deste trabalho é o estudo das condições necessárias e suficientes de otimalidade para problemas de controle ótimo, compreendendo o estudo do Princípio do Máximo e convexidade generalizada.

As condições necessárias dadas pelo Princípio do Máximo Pontryagin são obtidas mediante o formalismo de Dubovitski e Milyutin, que permite determinar, usando a linguagem da análise funcional, condições necessárias de otimalidade para uma classe de problemas extremos.

As condições suficientes de otimalidade são dadas introduzindo a noção de invexidade generalizada adequadas ao problema, que denominaremos PM-pseudo-invexidade.

Palavras-chave: Condições de Otimalidade, Formalismo de Dubovitskii-Milyutin, Princípio do Máximo, Invexidade Generalizada.

ABSTRACT

The purpose of this work is the study of necessary and sufficient conditions of optimality for optimal control problem including the study of the Maximum Principle and generalized convexity.

The necessary conditions given by the Pontryagin Maximum Principle are obtained by means of the Dubovitski and Milyutin formalism, which allows to determine, using the language of functional analysis, necessary optimality conditions for a class of extreme problems.

The sufficient conditions of optimality are given by introducing the notion of generalized invexity suitable to the problem, which we will call PM-pseudo-invexity.

Keywords: Optimality Conditions, Dubovitskii and Milyutin formalism, Maximum Principle, Generalized Invexity.

SUMÁRIO

CAPÍTULO 1

INTRODUÇÃO

Considera-se problemas de extremo a todo aquele problema de minimizar ou maximizar um funcional sob um dado conjunto de pontos viáveis, que comumente denominamos Problema de Otimização.

A teoria de Otimização, seja diferenciável ou não diferenciável é baseado principalmente na procura de condições necessárias e suficientes que permitem caraterizar soluções ótimas. As condições necessárias se fundamentam na hipótese de que o conjunto de soluções possíveis para o problema a resolver, verifica certas condições, chamadas restrições, que caraterizam os cones tangente e/ou cones factíveis através de derivadas de primeira ordem dos funcionais que definem as restrições (cone tangente para as restrições de igualdade e cones factíveis para o caso das restrições de desigualdade). Problemas em que esta hipótese é verificado são chamados de problemas regulares. Entretanto, existem problemas onde essa hipóteses não é garantida, esse tipo de problemas são chamados de problemas não regulares ou degenerados. Consequentemente, as condições de otimalidade de primeira ordem de Karush-Kuhn-Tucker, não são aplicáveis. Portanto é necessário estabelecer condições de otimalidade baseado em aproximações ao conjunto factível que não sejam de primeiro ordem, ou seja, usando derivadas de ordem superior, que de fato coincidem com as de primeira ordem quando o problema é regular, ver Orellana [42]. Quando a finalidade do problema é minimizar (ou maximizar) que depende de variáveis, que a sua vez são funções, as quais variam no tempo, então estamos em um problema de otimização dinâmica. Esses problemas caraterizam-se pelo fato que o funcional a minimizar (ou maximizar) é um operador, isto é, é uma função definida sobre um espaço de funções, de dimensão infinita.

Esta classe de problemas incluem os problemas variacionais, onde os funcionais a minimizar (ou maximizar) é um operador integral, cuja origem se atribui ao planteamento do problema da Braquistócrona, formulado por Bernoulli em 1686. O problema do Braquistocróna consiste em encontrar a trajetória ótima de uma partícula que, sujeita a um campo gravitacional constante, sem atrito e com velocidade inicial nula, se desloca entre dois pontos no menor intervalo de tempo, ou seja, a curva de descenso mais rápido. A teoria deste tipo de problemas é chamada de Cálculo de Variações.

O resultado principal do cálculo de variações que permite determinar curvas extremais são as equações de Euler e Lagrange, introduzidas por Euler (1707-1783) e Lagrange (1736-1813). Estas equações estabelecem condições necessárias de otimalidade, ou seja, se uma curva ou trajetória é um extremo, mínimo ou máximo de um operador que determina um problema variacional, então necessariamente deve satisfazer as equações de Euler-Lagrange.

Uma outra classe importante de problemas de otimização dinâmica são os problemas de controle ótimo, que na verdade são uma generalização natural do cálculo de variações, e tem muitas aplicações, na medicina, ver Pinch [36], na física, ver Christensen [11], na biologia, ver Swam [40], e em particular na economia, modelos neoclássicos de crescimento econômico, ver Macki [32], Shone [39], Cerda [9] entre outras.

O objetivo deste trabalho é estudar condições necessárias e suficientes de otimalidade para problemas de controle ótimo, para isso faremos uma pequena introdução no que se refere à definição, formulação e resultados principais. Um problema de controle ótimo consiste basicamente em maximizar (ou minimizar) um funcional, sujeito a um sistema que evoluciona no tempo (sistema dinâmico). Este funcional depende de uma variável, que descreve o estado do sistema em cada momento, e uma variável de controle, que regula o comportamento do sistema. Estas variáveis se relacionam por uma equação diferencial ordinária (ou em derivadas parciais) que chamamos de sistema dinâmico.

Pensemos em um sistema dinâmico que evoluciona no tempo (como por exemplo o corpo humano ou um sistema econômico), com um horizonte temporal predefinido $[t_0, T]$, onde a situação inicial é dado, digamos x_0, e que a evolução do sistema pode ser influenciada por decisões de um agente. O estado do sistema é descrito, em cada momento, por uma variável $x(t)$, denominada ***variável de estado***, e as decisões consideradas pelo agente são representadas por uma variável $u(t)$, denominada ***variável de controle***. O controle $u(t)$ exerce uma influencia sobre o comportamento do sistema dinâmico, descrita pela variável de estado $x(t)$, em qualquer instante $t \in [t_0, T]$. Portanto a evolução do sistema depende das decisões tomadas pelo agente, representada por $u(t)$.

As variáveis de estado $x(t)$ e a variável de controle $u(t)$ se relacionam pelo sistema de equações diferencias, denominadas ***equação de estado***,

$$\begin{aligned} x'(t) &= \varphi(x(t), u(t), t),\ \ t \in [t_0, T], \\ x(t_0) &= x_0, \end{aligned}$$

onde $x : [t_0, T] \longrightarrow \mathbb{R}^n$, $u : [t_0, T] \longrightarrow \mathbb{R}^r$ e $\varphi : \mathbb{R}^n \times \mathbb{R}^r \times [t_0, T] \longrightarrow \mathbb{R}^n$.

A variável de controle $u(\cdot)$ usualmente pertence a uma família de funções U chamado ***conjunto de funções admissíveis***, com valores em um conjunto $\mathbb{U}$ de $\mathbb{R}^r$, que representa as restrições físicas do controle $u(t)$ em um tempo t. O conjunto $\tilde{U} = \{u(t) : u \in U,\ t \in [t_0, T]\}$, de imagens de controles admissíveis, se conhece como ***região de controle***. A família de funções U e a função $\varphi(x, u, t)$ devem satisfazer hipóteses suficientes, para garantir a existência e unicidade de soluções da ***equação de estado***.

Uma vez escolhido um controle $u(\cdot)$ em U, a equação de estado determina uma trajetória ou estado $x(t)$, com $x(t_0) = x_0$. A solução da equação de estado, dependerá das diferentes escolhas do valor da variável de controle, isto é, escolhas diferentes do controle implica em trajetórias diferentes do sistema dinâmico.

Uma das variantes mais simples de um problema de controle ótimo é a seguinte: Dado um sistema dinâmico que evoluciona no tempo, de acordo com a equação de estado, se deve obter as variáveis, de estado $x(t)$ e de controle $u(t)$ factíveis, isto é, funções $x(t)$ e $u(t)$ que satisfazem a equação de estado e que minimizem (ou maximizem) o operador integral

$$F(x, u) = \int_{t_0}^{T} \Phi(x(t), u(t), t)dt + S(x(T)), \tag{1.1}$$

onde, a variável de controle satisfaz a restrição do tipo

$$u(t) \in \mathbb{U},\ \text{para quase todo } t \in [t_0, T],$$

com $\mathbb{U}$ subconjunto de $\mathbb{R}^r$, o tempo T é dado, onde supomos o estado no tempo final, livre, isto é, $x(T)$ é variável e $\Phi : \mathbb{R}^n \times \mathbb{R}^r \times [t_0, T] \longrightarrow \mathbb{R}$.

O funcional $\Phi(x, u, t)$ medi quantitativamente o comportamento do sistema. A integral de $\Phi(x, u, t)$ dada por (1.1) depende das variáveis de estado $x(t)$ e de controle $u(t)$ ao longo de um horizonte temporal, ou seja, ao longo de um intervalo de tempo $[t_0, T]$, portanto, representa o valor do comportamento do sistema através do tempo, enquanto a função $S(x(T))$ representa a valorização do estado que fica o sistema ao final do horizonte temporal.

Existem muitas variantes dos problemas de controle ótimo que diferem da formulação inicial, no que refere às formas que pode tomar o funcional objetivo, condições iniciais e finais de estado e as restrições sob o controle e/ou a variável de estado. Em alguns problemas pode ser que se tenha $x(T)$ fixo, ou que seja limitado superior ou inferiormente, ou que o horizonte temporal não seja limitado por um valor dado T. No problema anterior não são considerados restrições nas variáveis de estado $x(t)$, embora em alguns problemas é necessário impor restrições de igualdade ou desigualdade à variável de estado $x(t)$.

Existem diferentes formas para o funcional objetivo (1.1), dependendo das funções $\Phi(x,u,t)$ e $S(x(T))$. São denominados: ***Bolza***, se $\Phi(x,u,t)$ e $S(x(T))$ são não nulos; ***Mayer***, se $\Phi(x,u,t)$ é nulo e $S(x(T))$ não nulo e ***Lagrange***, se $S(x(T))$ é nulo e $\Phi(x,u,t)$ não nulo. Mostra-se que los problemas de controle ótimo com restrições do tipo Mayer, Bolza e Lagrange são equivalentes, ver Cerda [9]. Neste trabalho consideraremos problemas de controle ótimo com restrições do tipo Lagrange.

Um problema de controle ótimo com restrições do tipo Lagrange, onde não existe restrições no controle ($\mathbb{U} = \mathbb{R}^r$) é chamado de problema de controle de Lagrange, neste caso o problema é

$$\min \int_{t_0}^{T} \Phi(x(t), u(t), t)dt$$
$$\text{s.a } x'(t) = \varphi(x,u,t),\ x(t_0) = c,\ x(T) = d.$$

O objetivo da teoria de controle é o estudo das condições necessárias e suficientes de otimalidade, além de resultados para garantir a existência e unicidade da solução. Assim como também do desenvolvimento de métodos para sua solução desde o ponto de vista teórico e aplicações.

Os principais resultados da teoria de controle ótimo são o Princípio de Máximo de Pontryagin (1962), e o princípio do Máximo Local de Pontryagin (equações de Euler-Lagrange). A diferencia principal entre esses dois princípios é que no princípio do máximo local, não se tem a maximização da Hamiltoniana em si, o que se tem é uma condição necessária para a maximização da Hamiltoniana. Em quanto no princípio do Máximo, o Hamiltoniano é maximizado.

O princípio do Máximo local estabelece o seguinte:

Seja (x^0, u^0) uma solução do problema de controle de Lagrange:

$$\begin{cases} \min F(x,u) = \displaystyle\int_{t_0}^{T} \Phi(x(t), u(t), t)dt \\ s.a \\ \qquad x'(t) = \varphi(x(t), u(t), t) \ x(t_0) = c, \ x(T) = d \\ \qquad x(t) \in \mathbb{U} \text{ q.t.p. } t_0 \leq t \leq T, \end{cases} \tag{1.2}$$

onde as funções $\Phi : \mathbb{R}^n \times \mathbb{R}^r \times [t_0, T] \longrightarrow \mathbb{R}$ e $\varphi : \mathbb{R}^n \times \mathbb{R}^r \times [t_0, T] \longrightarrow \mathbb{R}^n$ são funções continuamente diferenciáveis com respeito a (x, u) e mensuráveis em t, onde $\mathbb{U}$ é um conjunto fechado e convexo, com interior não vazio de $\mathbb{R}^r$. Então existem $\lambda_0 \geq 0$, e uma função $\psi : [t_0, T] \longrightarrow \mathbb{R}^n$ simultaneamente não nulas, tais que

$$\psi'(t) = -\varphi_x^*(x^0, u^0, t)\psi(t) + \lambda_0 \Phi_x(x^0, u^0, t); \tag{1.3}$$

além disso, para todo $u(t) \in \mathbb{U}$,

$$\left\langle \lambda_0 \Phi_u(x^0, u^0, t) - \varphi_x^*(x^0, u^0, t)\psi(t), u - u^0 \right\rangle \geq 0 \text{ q.t.p. } t_0 \leq t \leq T, \tag{1.4}$$

onde $\varphi_x^*(x^0, u^0, t)$ e $\varphi_u^*(x^0, u^0, t)$ denotam o operador adjunto de $\varphi(x^0, u^0, t)$ e $\varphi(x^0, u^0, t)$ respetivamente.

O princípio do Máximo Local, e suas variantes proporcionam condições necessárias de otimalidade para distintas formulações do problema de controle, ver Alekseev [1], Cerda [9], Girsanov [22], Craven [14] sempre sob a hipótese No controle u tal que $u(t) \in \mathbb{U}$ é um conjunto fechado e convexo de $\mathbb{R}^r$. No entanto em muitas aplicações é natural considerar problemas onde a variável de controle $u(t)$ pertence a um conjunto menos restritivo, como são os problemas onde o controle ótimo toma somente dois valores, por exemplo $u^0(t) = a$ ou $u^0(t) = b$, neste caso o controle u^0 é denominado controle Bang-Bang, ver Cerda [9] e Craven [14]. Este tipo de fenômenos acontece quando o funcional objetivo e a equação de estado é linear na variável de controle $u(t)$.

A derivação de condições de otimalidade para problemas de controle ótimo, sob o suposto que a variável de controle $u(t)$ pertencença a um conjunto arbitrário $\mathbb{U}$ de $\mathbb{R}^r$ são estabelecidos pelo Principio de Máximo de Pontryagin e suas variantes, que estudaremos no Capítulo 5.

Ambos princípios são aplicáveis somente no caso em que o problema em estudo seja regular, o que em teoria de controle é denominado: Problema controlável ou extremais regulares ou normais, isto é, quando o multiplicador associado ao funcional objetivo, é não nulo, ou seja λ_0 é não nulo. As condições suficientes pra garantir $\lambda_0 > 0$ são denominadas condições de regularidade ou ou normalidade, ver Girsanov [22] e Craven [14] e no contexto da otimização matemática, são denominados restrições de qualificação, ver Bazaara [3]. Em caso λ_0 seja nulo, dizemos que o problema é não regular, ver Avakov [25], neste caso o Principio do Máximo Local e o Principio de Máximo de Pontryagin proporcionam condições degeneradas, não dependem do funcional a minimizar e não produzem nenhuma informação sobre o extremo local ou global em estudo. A regularidade é fundamental, tanto para a obter as condições necessárias de otimalidade não degeneradas, como para a obtenção das condições suficientes.

Existem várias formas de provar o Princípio do Máximo local de Pontryagin, uma de elas é via analise real, ver Pontryagin [37], outra via inclusões diferenciais, ver Challco [10], via métodos variacionais, veja Cerda [9] e também usando princípio de Lagrange em dimensão infinita, veja Ioffe [24], [26] ou através do formalismo de Dubovitskii e Milyutin, ver Girsanov [22], Orellana [42]. Esta última forma será vista com detalhe no Capítulo 4.

Para formular condições necessárias de otimalidade para problemas extremos, mediante este formalismo se deve caraterizar, o cone das direções factíveis ao conjunto das restrições de desigualdade, cone das direções tangentes, ao conjunto das restrições de igualdade num ponto ótimo e seus respetivos cones duais, ver Girsanov [22].

Se as funções involucradas no problema são convexas, o qual determina um problema convexo, então as condições necessárias de otimalidade são também suficientes, mas estas hipóteses são muitos restritivas, dado que existem uma classe amplas de problemas não convexos. Com a finalidade enfraquecer esta hipótese de convexidade surgiu na literatura a noção de ***função invexa*** no contexto da otimização matemática, introduzida por Hanson em 1981, em [23] e suas generalizações em Craven [13] e Martin [33] .

Em 1985, Martin em [33] introduz uma generalização da noção invexidade, denominada KT-invexidade e mostrou que KT-invexidade não é unicamente uma condição suficiente de otimalidade, para problemas de otimização clássicas, no sentido de que cada ponto satisfaz as condições de Kuhn-Tucker, (definido por Craven em [13], como KT-ponto) é um mínimo global do problema, se não também uma condição necessária, isto é, mostra que cada KT-ponto é um mínimo global se, e somente se, o problema é KT-invexo. A classe mais ampla de problemas que cumpre esta proposição são a classe de problemas KT-invexos. Posteriormente esta noção foi aplicada a problemas em tempo contínuo por Brandão em [6] e [38], além disso, para certo tipo de problemas variacionais por Mond e Hussain em [35], anos depois Oliveira et al. em [18] estenderam a noção de KT-invexidade da programação matemática para problemas de controle ótimo regulares, para mostrar que as condições necessárias de otimalidade, dadas por uma versão do Princípio de Máximo de Pontryagin são suficientes se e somente se, o problema é KT-invexo. Tambem foram extendidas recentemente para problemas de controle no suave e multiobjetivo em [15], [16] e [17].

Este trabalho está organizado da seguinte maneira. No Capítulo 2 apresentamos as definições e resultados essenciais para a determinação de condições necessárias de extremo para problemas extremais, e determinação do formalismo de Dubovitskii e Milyutin.

No Capítulo 3 estudamos o formalismo de Dubovitskii e Milyutin. Na Seção 3.1 definiremos conjuntos associados às restrições de certo problema extremo, como o conjunto de direções de factíveis e tangentes, assim como também os conjuntos de direções de descida do funcional objetivo, para logo mostrar um resultado que dá condições necessárias e suficientes de otimalidade para um problema extremal com restrições de igualdade e desigualdade. Na seção 3.2 veremos o estudo dos cones duais associados aos conjuntos de direções de descida, factível e tangente, assim como alguns resultados que caraterizam esses conjuntos.

No Capítulo 4 estudaremos uma aplicação do formalismo de Bubovitskii e Milyutin, demonstrar o Princípio de Máximo Local de Pontryagin usando o formalismo de Bubovitskii e Milyutin.

Finalmente, no Capítulo 6 são desenvolvidas as condições suficientes de otimalidade para os problemas de controle estudados nos Capítulos 4 e 5. Na primeira seção são abordados a suficiência sob hipóteses de convexidade, e na segunda seção introduzimos a noção de MP-pseudo-invexidade para mostrar a condição suficiente de otimalidade sob hipoteses de MP-pseudo-invexidade.

CAPÍTULO 2

PRELIMINARES E RESULTADOS BÁSICOS

Neste capítulo iremos definir alguns conceitos de análise funcional e análise convexa, assim como alguns resultados importantes que serão de muita utilidade no desenvolvimento desta dissertação. As definições e resultados desse capítulo podem ser encontrados em Boyd [5], Brézis [7], Kreyszig [28] e Girsanov [22].

2.1 Conjuntos Convexos, Cones e Cones Duais

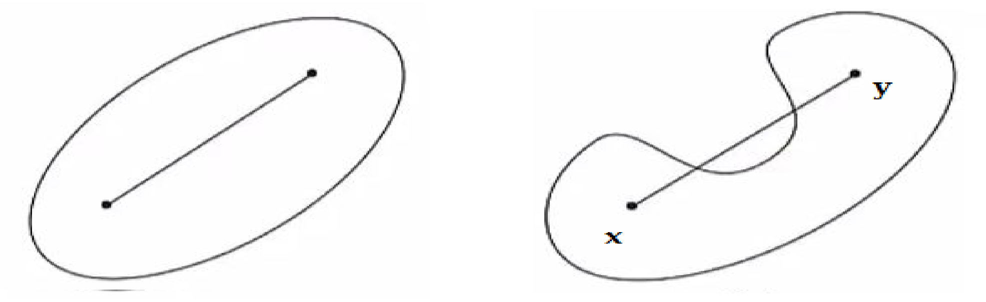
Nesta seção consideramos E conjunto não-vazio, a menos de menção contrária, dotamos de certa estrutura ou propriedade.

Definição 2.1. *Um conjunto $C \subset E$ é dito **convexo** quando dados $x, y \in C$ o segmento*

$$[x\ y] = \{(1-t)x + ty : t \in [0; 1]\} \tag{2.1}$$

estiver inteiramente contido em C.

A Figura 2.1 ilustra dois conjuntos, um convexo e outro não.



(a) conjunto convexo (b) conjunto não convexo

Figura 2.1: Conjunto convexo e não convexo

Uma classe de funções que tem propriedades importantes no contexto de otimização são as funções convexas.

Definição 2.2. *Seja $C \subset E$ um conjunto convexo. Dizemos que a função $f : E \to \mathbb{R}$ é **convexa** em C se*

$$f((1-t)x + ty) \leq (1-t)f(x) + tf(y) \tag{2.2}$$

para todos $x, y \in C$ e $t \in [0, 1]$.

Geometricamente, significa que o gráfico está sempre abaixo do segmento de reta que liga as extremidades, como na Figura abaixo.

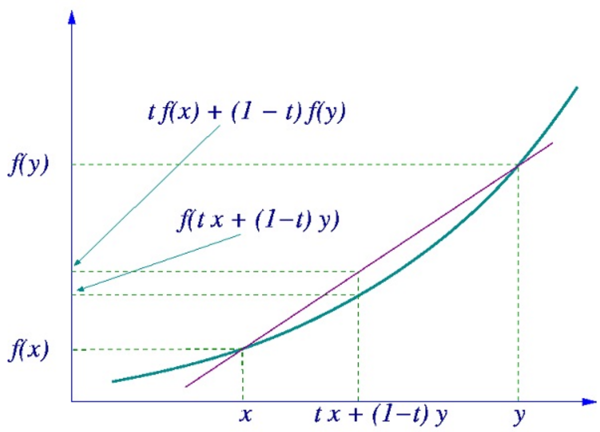


Figura 2.2: Função convexa

Definição 2.3. *Um conjunto $K \subset E$, com E espaço linear, é dito ser um **cone com vértice** em 0, se para cada $x \in K$, $\lambda x \in K$ para qualquer $\lambda > 0$, isto é, $\lambda K = K$.*

Observação 2.1. *Se K é um cone com vértice em 0, então x_0+K é um cone com vértice em x_0.*

Teorema 2.1. *Seja K um cone com vértice em x_0 e $f(x)$ um funcional linear sobre E tal que $f(x) \geq \alpha$, para $x \in K$. Então $f(x) \geq f(x_0)$ $\forall x \in K$.*

Demonstração. Suponha, por absurdo, que existe $y \in K$ e $f(y) < f(x_0)$. Por definição de cone, $x_0+t(y-x_0)$, $t > 0$, deve estar inteiramente contido em K, mas f é linear, portanto

$$f(x_0 + t(y - x_0)) = f(x_0) + t(f(y) - f(x_0)).$$

Como $f(y) < f(x_0)$, então $f(x_0 + t(y - x_0)) \to -\infty$, para t suficientemente grande. Isto contradiz o fato de $f(x) \geq \alpha$ $\forall x \in K$. □

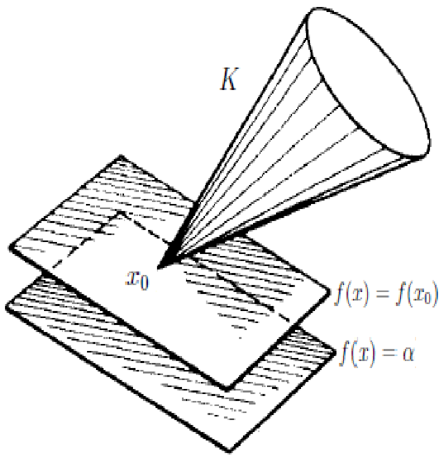


Figura 2.3: Interpretação geométrica do Teorema (2.1)

Agora introduziremos o conceito importante de cone dual que será de muita importância na caraterização das condições necessárias de otimalidade.

Definição 2.4. *Seja K um cone em E, com vértice em 0. O conjunto K^* de todos os funcionais $f \in E'$ tais que $f(x) \geq 0$ para todo $x \in K$ é chamado de* ***cone dual*** *de K:*

$$K^* = \{f \in E' : f(x) \geq 0 \ \forall x \in K\}. \tag{2.3}$$

Denotamos por E' o dual topológico do espaço de Banach E, isto é, E' é o conjunto de dos funcionais lineares e contínuos, definidos sobre E. É fácil mostrar que E' é um espaço de Banach com a norma

$$\|f\| = \sup\{|f(x)| : x \in E, \|x\| \leq 1\}.$$

Exemplo 2.1. *Sejam $E = \mathbb{R}^n$ e $K = \mathbb{R}^n_+ = \{(x_1, x_2, .., x_n) : x_1, ..., x_n > 0\}$. Então K é um cone fechado em E, além disso $K^* = K$.*

Exemplo 2.2. *Sejam $E = \mathbb{R}^2$ e o cone K associado a E dado por*

$$K = \left\{x = (x_1, x_2) : a^T x \geq 0, b^T x \geq 0; \textit{com } a, b \in \mathbb{R}^2 \textit{ fixos}\right\}.$$

Então

$$K^* = \left\{f \in \mathbb{R}^2 : \lambda_1 a + \lambda_2 b;\ \ \lambda_1, \lambda_2 \geq 0\right\}.$$

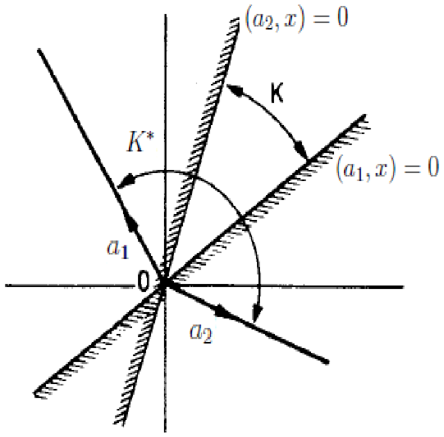


Figura 2.4: Representação gráfica de K^* do Exemplo 2.2

Observação 2.2. *Se K é um cone em E com vértice em 0, temos que*

i) *K^* é um cone de vértice 0,*

ii) *Se $K = \{0\}$, então $K^* = E'$,*

iii) *Se K é um subespaço arbitrário de um espaço de Banach E, então o cone dual associado a ele é*

$$K^* = \{f \in E' : f(x) = 0\ \forall x \in K\}.$$

Neste casso K^* é chamado de subespaço anulador ou aniquilador do subespaço K.

Lema 2.1. *Sejam K_1 e K_2 cones com vértice na origem em E. Se $K_1 \subset K_2$, então $K_2^* \subset K_1^*$.*

Demonstração. Se $f \in K_2^*$ então $f(x) \geq 0\ \forall x \in K_2 \supset K_1$. Então $f(x) \geq 0\ \forall x \in K_1$. Isto implica que $f \in K_1^*$. $\square$

Lema 2.2. *Seja $K_{\alpha\in\Lambda}$ uma família de cones, então* $\left(\bigcup_{\alpha\in\Lambda} K_\alpha\right)^* = \bigcap_{\alpha\in\Lambda} K_\alpha^*$ *para qualquer conjunto de índice* Λ.

Demonstração. Se $f \in \left(\bigcup_{\alpha\in\Lambda} K_\alpha\right)^*$ então $f(x) \geq 0\ \forall x \in \bigcup_{\alpha\in\Lambda} K_\alpha$. Fixando $\alpha \in \Lambda$, temos que $f(x) \geq 0\ \forall x \in K_\alpha$, e então $f \in K_\alpha^*$. Como $\alpha \in \Lambda$ é arbitrário, segue que $f \in \bigcap_{\alpha\in\Lambda} K_\alpha^*$.

Por outro lado, se $f \in \bigcap_{\alpha\in\Lambda} K_\alpha^*$, temos que $f \in K_\alpha^*\ \forall\alpha \in \Lambda$, ou seja, $f(x) \geq 0\ \forall x \in K_\alpha \subset \bigcup_{\alpha\in\Lambda} K_\alpha$ e $\forall\alpha \in \Lambda$. Portanto, $f(x) \geq 0$ para todo $x \in \bigcup_{\alpha\in\Lambda} K_\alpha$, donde $f \in \left(\bigcup_{\alpha\in\Lambda} K_\alpha\right)^*$. □

Lema 2.3. *Sejam K_i cones convexos para $i = 1, \dots, n$. Então* $\bigcup_{i=1}^{n} K_i^* = \sum_{i=1}^{n} K_i^*$.

Demonstração. Veja Girsanov [[22], p. 34]. □

Lema 2.4 (Teorema de Krein)**.** *Sejam $\hat{K}$ um cone convexo com vértice na origem 0, o qual contém pontos interiores, e $\hat{L}$ um subespaço tal que int $(\hat{K}) \cap \hat{L} \neq \emptyset$. Seja $\bar{g}(x)$ uma forma linear sobre $\hat{L}$ tal que $\bar{g}(x) \geq 0$ sobre $\hat{K} \cap \hat{L}$. Então existe uma forma linear contínua $g(x)$ sobre o espaço de Banach E, de tal maneira que*

$$\begin{cases} g(x) = \bar{g}(x) \text{ para todo } x \in \hat{L}; \\ g(x) \geq 0 \text{ para todo } x \in \hat{K}. \end{cases}$$

Demonstração. A demonstração de este resultado pode ser encontrada em Girsanov [[22] p.31]. □

Teorema 2.2. *Sejam $K_1, \dots, K_n$ cones convexos abertos em E, tais que*

$$\bigcap_{i=1}^{n} K_i \neq \emptyset,$$

então

$$\left(\bigcap_{i=1}^{n} K_i\right)^* = \sum_{i=1}^{n} K_i^*.$$

Demonstração. Usaremos argumentos geométricos tais como conceitos de produto direto de espaços de Banach e do respectivo espaço dual.

Definamos o espaço $M = \prod_{i=1}^{n} E$, onde E é Banach, isto é, o espaço

$$M = \{\tilde{x} = (x_1, \dots, x_n) : x_i \in E,\ i = 1, \dots, n\}.$$

Assim

$$M^* = \{F = (f_1, \ldots, f_n) : f_i \in E^*,\ i = 1, \ldots, n\}.$$

Deste modo, podemos dizer que cada $F \in M^*$ tem a representação

$$F(\tilde{x}) = \sum_{i=1}^{n} f_i(x_i),\ \text{para todo } \tilde{x} = (x_1, \ldots, x_n), \tag{2.4}$$

onde $f_i \in E^*$, para todo $i = 1, \ldots, n$. Consideremos ainda, o conjunto

$$K = \{\tilde{x} = (x_1, \ldots, x_n) : x_i \in K_i,\ i = 1, \ldots, n\} \subset M.$$

Aqui, K é um cone convexo aberto, porque é o produto direto de cones convexos abertos ($K = \prod_{i=1}^{n} K_i$). Definamos $L = \{\tilde{x} = (x, \ldots, x) : x \in E\} \subset M$, onde L é um subespaço de M. Por hipótese, existe $x_0 \in K_i,\ i = 1 \ldots, n$, logo $L \cap K \neq \emptyset$.

Seja $f \in \left(\bigcap_{i=1}^{n} K_i\right)^*$. Tomemos o funcional linear $\bar{f}$ definido sobre L, como segue:

$$\bar{f}(\tilde{x}) = f(x), \tag{2.5}$$

onde $\tilde{x} = (x, \ldots, x) \in L$. Se $\tilde{x}_0 = (x_0, \ldots, x_0) \in L \cap K$, então $x_0 \in \bigcap_{i=1}^{n} K_i$. Assim, $f(x_0) \geq 0$. Portanto,

$$\bar{f}(\tilde{x}_0) \geq 0, \forall \tilde{x}_0 \in L \cap K.$$

Usando o fato de *int* $(K) = K$, pois K é aberto, temos que $L \cap int\ (K) \neq \emptyset$. Pelo Lema 2.4, conseguimos a existência de $F = (f_1, \ldots, f_n) \in M^*$, tal que

$$F(\tilde{x}) \geq 0\ \forall \tilde{x} \in K; \tag{2.6}$$

$$F(\tilde{x}) = \bar{f}(\tilde{x})\ \forall \tilde{x} \in L. \tag{2.7}$$

Assim, para todo $\tilde{x}_0 \in L$, temos por (2.4) que $F(\tilde{x}_0) = \sum_{i=1}^{n} f_i(x_0)$; por (2.7) $F(\tilde{x}_0) = \bar{f}(\tilde{x}_0)$ e por (2.5) $\bar{f}(\tilde{x}_0) = f(x_0)$. Logo,

$$f(x_0) = \sum_{i=1}^{n} f_i(x_0),\ \forall \tilde{x}_0 \in L.$$

Então, para $\tilde{x} = (x_0, \ldots, x_0)$; $x_0 \in \bigcap_{i=1}^{n} K_i$, também teremos

$$f(x_0) = \sum_{i=1}^{n} f_i(x_0), \text{ o que equivale a}$$

$$f = \sum_{i=1}^{n} f_i, \text{ onde } f \in \left(\bigcap_{i=1}^{n} K_i\right)^*.$$

Devido a (2.6) segue que $F(\tilde{x}) \geq 0$, onde $F(\tilde{x}) = \sum_{i=1}^{n} f_i(x_i)$, para todo $x_i \in K_i$, $i = 1, \ldots, n$ e $\tilde{x} = (x_1, \ldots, x_n)$. Consequentemente, $f_i(x_i) \geq 0$, $\forall x_i \in K_i$, ou seja, $f_i \in K_i^*$, $i = 1, \ldots, n$.

Agora de acordo com a análise feita para $\tilde{x}_0$ onde $x_0 \in \bigcap_{i=1}^{n} K_i$, podemos usar o resultado obtido acima e concluirmos que $f(x_0) = \sum_{i=1}^{n} f_i(x_0)$, com $f_i \in K_i^*$, $\forall i = 1, \ldots, n$. Portanto, dado $f \in \left(\bigcap_{i=1}^{n} K_i\right)^*$, existem $f_i \in K_i^*$, $i = 1, \ldots, n$, tais que $f = \sum_{i=1}^{n} f_i$. Em outras palavras, $\left(\bigcap_{i=1}^{n} K_i\right)^* \subset \sum_{i=1}^{n} K_i^*$.

Por outro lado, seja $f \in \sum_{i=1}^{n} K_i^*$, existem $f_i \in K^*$ tais que $f = \sum_{i=1}^{n} f_i$, para todo $i = 1, \ldots, n$. Se $x_0 \in \bigcap_{i=1}^{n} K_i$, temos que $f_i(x_0) \geq 0$ $\forall i = 1 \ldots, n$. Logo, $f(x_0) \geq 0$ $\forall x_0 \in \bigcap_{i=1}^{n} K_i$. Assim, $f \in (\bigcap_{i=1}^{n} K_i)^*$, ou seja, $\sum_{i=1}^{n} K_i^* \subset \left(\bigcap_{i=1}^{n} K_i\right)^*$. $\square$

2.2 Topologia Fraca e Teoremas de Separação

Nesta seção vamos definir alguns conceitos de topologia, e um resultado fundamental de análise funcional com muitas aplicações em análise convexa, o Teorema de Separação de Hanh Banach. Para mais detalhes veja Brézis [7], Kreyszig [28].

Definição 2.5. *Seja C um subconjunto não vazio de um espaço vetorial normado E.*

*i) Se C é fechado na topologia da norma (isto é, se dada uma sequência $(x_n) \subset E$ tal que $\|x_n - x\|_E \to 0$ quando $n \to \infty$, temos que $x \in C$), então dizemos que C é **fortemente fechado**.*

*ii) Dizemos que C é **fracamente fechado**, se ele for fechado na topologia fraca, ou seja,* *se dada uma sequência* $(x_n) \subset C$, *tal que* $f(x_n) \longrightarrow f(x)$ *quando* $n \longrightarrow \infty$ $\forall f \in E^*$, *então* $x \in C$.

iii) Dada uma sequência $(x_n) \subset C$. *Se* $f(x_n) \longrightarrow f(x)$ *quando* $n \longrightarrow \infty$ *para todo* $f \in J(x) \subset E^{**}$, *onde* $J(x)$ *é isomorfo a* E, *então* $x \in C$. *Diz-se que* C *é fechado na topologia fraca**, *isto é,* C *é* ***fracamente*** * ***fechado***

Observação 2.3.

i) Em espaços normados de dimensão finita as topologias da norma, fraca e fraca* *são equivalentes,*

ii) A topologia fraca* *é mais fraca que a topologia fraca de* E^* *e portanto qualquer conjunto fracamente*$*$ *fechado em* E^* *é fracamente fechado; estas topologias são equivalentes quando* E *é um espaço de Banach reflexivo.*

Teorema 2.3. *Seja* C *um conjunto convexo em* E. *Então* C *é fortemente fechado se, e somente se, ele é fracamente fechado*

Demonstração. Veja Brézis [[7] p.60]. □

Lema 2.5. *Se* K *é fracamente fechado então* $K^{**} = K$

Demonstração. Veja Girsanov [[22] p.34-35]. □

Teorema 2.4. *Sejam* K_i, $i = 1, \ldots, n$, *cones convexos e fracamente fechados tais que* $\sum_{i=1}^n K_i^*$ *é fracamente** *fechado, então*

$$\left(\bigcap_{i=1}^n K_i\right)^* = \sum_{i=1}^n K_i^*.$$

Demonstração. Seja $Q = \sum\limits_{i=1}^n K_i^*$, então

$$\begin{aligned} Q^* &= \left(\sum_{i=1}^n K_i^*\right)^* = \left(\bigcup_{i=1}^n K_i^*\right)^* \\ &= \bigcap_{i=1}^n K_i^{**} \\ &= \bigcap_{i=1}^n K_i \end{aligned}$$

Assim $Q^{**} = \left(\bigcap_{i=1}^n K_i\right)^*$. Mas $Q^{**} = Q$ e portanto segue o resultado. □

Teorema 2.5. *Seja E um espaço de Banach de dimensão finita. Então:*

i) K^* *é fechado.*

ii) $K^* = (clK)^*$, *ou seja o cone dual de um cone e de seu fecho são iguais.*

Demonstração.

i) Seja $(f_n) \subset K^*$, com $\lim_{n\to\infty} f_n = f$. Fixando $x \in K$ arbitrário e passando o limite quando $n \to \infty$ em $f_n(x) \geq 0$, obtemos $f(x) \geq 0$. Portanto, como $x \in K$ é arbitrário, $f \in K^*$ isto mostra que K^* é fechado.

ii) Temos que $K \subseteq \text{cl}K$, logo pelo Lema 2.1 $\text{cl}K^* \subset K^*$. Por outro lado, seja $f \in K^*$ e $x \in \text{cl}K$ arbitrário. Então existe uma sequência $(x_n) \subset K$ tal que $\lim_{n\to\infty} x_n = x$. Agora passando o limite quando $n \to \infty$ em $f(x_n) \geq 0$, como f é contínuo, obtemos $f(x) \geq 0$. Portanto $f \in (clK)^*$, isto completa a prova.

□

Definição 2.6. *Um **hiperplano afim** é um subconjunto H de E da forma*

$$H = \{x \in E : f(x) = \alpha\},$$

onde $f(x)$ é um funcional linear que não nulo e $\alpha \in \mathbb{R}$ é uma constante dada. Escrevemos $H = [f = \alpha]$ e dizemos que $f = \alpha$ é a equação de H.

Proposição 2.1. *O hiperplano $H = [f = \alpha]$ é fechado se, e somente se, f é contínua.*

Demonstração. Veja Bréziz [[7] p.5]. □

Definição 2.7. *Sejam $A, B \subset E$ não vazios. Dizemos que um funcional linear contínuo não-nulo f **separa** os conjuntos A e B (ou que o hiperplano $H = [f = \alpha]$ separa A e B), se existe $\alpha \in \mathbb{R}$ tal que*

$$f(x) \leq \alpha \leq f(y), \ \forall x \in A \ \forall y \in B.$$

*Dizemos que $f(x)$ **separa estritamente** A e B se existe algum $\varepsilon > 0$ tal que*

$$f(x) \leq \alpha - \varepsilon \ \forall x \in A \ e \ f(x) \geq \alpha + \varepsilon \ \forall x \in B.$$

Geometricamente, a separação significa que um conjunto encontra-se num dos semiplanos determinada por H, e o outro conjunto encontra-se no outro; veja a figura abaixo.

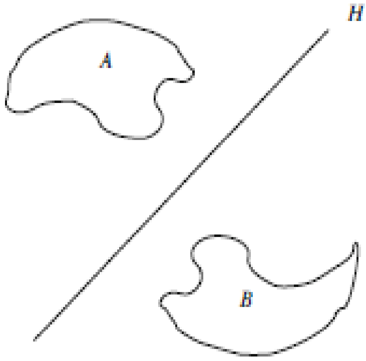


Figura 2.5: Separação de conjuntos

Teorema 2.6. *Se os conjuntos $A \subset E$ e $B \subset E$ são convexos, não-vazios e disjuntos, e A é aberto, então existe um funcional linear contínuo não-nulo separando A e B.*

Demonstração. Veja Brézis [[7] p.5]. □

Teorema 2.7. *Sejam $A \subset E$ e $B \subset E$ subconjuntos convexos, não-vazios e disjuntos. Se A é fechado e B compacto, então existe um funcional linear contínuo não-nulo separando A e B estritamente.*

Demonstração. Veja Brézis [[7] p.7]. □

Definição 2.8. *Dizemos que um funcional linear contínua $f(x)$ é **funcional suporte** para um conjunto $A \subset E$ em $x_0 \in A$ se,*

$$f(x) \geq f(x_0), \ \forall x \in A.$$

Sob as condições da definição acima, um hiperplano fechado $H = \{x : f(x) = f(x_0)\}$ é chamado **hiperplano suporte** para A em x_0.

O sentido geométrico de um **hiperplano suporte** é bastante simples: o conjunto A encontra-se em um lado do hiperplano e corta em um ponto, veja a Figura 2.6.

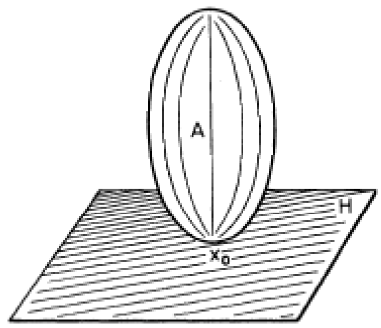


Figura 2.6: Hiperplano suporte

Observação 2.4.

- *Não existe funcional suporte alguma no interior de um conjunto.*
- *Se o conjunto A é um conjunto convexo, então A tem um funcional suporte em cada ponto da fronteira, veja Girsanov [[22] p.25].*

CAPÍTULO 3

FORMALISMO DE DUBOVITSKII-MILYUTIN

Uma grande quantidade de trabalhos na teoria de problemas extremos usam o formalismo de Dubovitskii-Milyutin, para obter condições necessárias de otimalidade (ver [22], [29], [30], [31] e entre outros).

Consideremos inicialmente um problema extremo, de minimizar um funcional definido sobre um aberto de um espaço de Banach sujeito a restrições, determinadas por conjuntos com interior vazio e não-vazio. Especificamente, seja E espaço de Banach, defina um funcional $f : U \subset E \longrightarrow \mathbb{R}$ e considere o seguinte problema extremo:

$$(P_0) \begin{cases} \min f(x) \\ s.a \\ x \in \bigcap_{i=1}^{n+1} Q_i, \ x \in U. \end{cases}$$

Na formulação usual, os conjuntos $Q_i, i = 1, ...n$, representam restrições de desigualdade e Q_{n+1} a restrição de igualdade do problema.

Na teoria apresentada por Dubovitskii e Milyutin em 1962, foram estabelecidas as condições necessárias de otimalidade local num ponto dado $x_0 \in E$, mediante um enfoque, em princípio, não geométrico, mas sim analítico, isto é, as condições necessárias de otimalidade são determinadas em forma de uma única equação, descrita mediante a linguagem da análise funcional. Elas são obtidas a partir da separação das aproximações cônicas aos conjuntos de restrições (de igualdade e/ou desigualdade) Q_i, $i = 1, .., n+1$, e ao conjunto $\{x \in E : f(x) < f(x_0)\}$. Neste estudo as definições foram feitas de tal modo que se os cones (de direções fatíveis, tangentes e de descida) são não vazias e convexos, então é válida a seguinte proposição. Se x_0 é um mínimo local do Problema (P_0), então não existe uma direção comum a todos os cones aproximantes. Os resultados obtidos por Dubovitskii e Milyutin provam que esta propriedade geométrica de otimalidade local num x_0 pode ser equivalentemente descrita em termos das formas lineares dos correspondentes cones duais.

Na primeira seção iremos a definir o conjunto direções de descida, factível e tangente, logo mostraremos um resultado devido a Dubovitski e Milyutin que dá uma condição necessária de otimalidade para o problema (P_0), o qual será uma ferramenta muito útil para mostrar o Princípio de Máximo Local de Pontryagin. Na segunda seção estudamos o cálculo de cones e seus duais para os conjuntos definidas na primeira seção.

3.1 Condições Necessárias de Extremo

A seguir vamos lembrar algumas definições e resultados necessários para abordar o formalismo de Dubovitskii-Miltyutin, para logo obter condições necessária de otimalidade para um problema de controle ótimo no capítulo seguinte. Para isso vamos seguir Alekseev [1] e Girsanov [22] .

Definição 3.1. *Dizemos que um vetor $h \in E$ é uma **direção de descida** do funcional $f(x)$ num ponto $x_0 \in E$, se existe uma vizinhança U de h, um escalar $\alpha = \alpha(f, h, x_0)$, $\alpha < 0$ e $\varepsilon_0 > 0$ tais que para todo $\varepsilon \in]0, \varepsilon_0[$ e qualquer $\bar{h} \in U$,*

$$f(x_0 + \varepsilon\bar{h}) \leq f(x_0) + \varepsilon\alpha.$$

Definição 3.2. *Um funcional $f(x)$ é **regularmente decrescente** ou **regularmente de descida**, se o conjunto das direções de descida no ponto x_0 é convexo. O conjunto de tais direções é denotado por $D(f; x_0)$ ou K_D.*

Lema 3.1. *As direções de descida geram um cone aberto $D(f; x_0)$ com vértice em 0.*

Demonstração. Seja $h \in E$ uma direção de descida do funcional f no ponto x_0. Então existem, uma vizinhança U de h, $\alpha < 0$ e $\varepsilon_0 > 0$, tais que para todo $\varepsilon \in]0, \varepsilon_0[$, tem-se

$$f(x_0 + \varepsilon\bar{h}) \leq f(x_0) + \varepsilon\alpha.$$

Seja $\lambda > 0$, tome uma outra vizinhança $V = \lambda U$ de h, $\widetilde{\varepsilon}_0 = \frac{\varepsilon_0}{\lambda}$ e $\widetilde{\alpha} = \lambda\alpha$. Logo para todo

$\widetilde{\varepsilon} \in]0, \widetilde{\varepsilon}_0[$ e qualquer $\widetilde{h} \in V$ temos

$$\begin{aligned} f(x_0 + \widetilde{\varepsilon}\widetilde{h}) &= f\left(x_0 + \frac{\varepsilon(\lambda\bar{h})}{\lambda}\right), \bar{h} \in U \\ &\leq f(x_0) + (\varepsilon\lambda)\alpha, \text{ (pois } \varepsilon < \widetilde{\varepsilon_0} \Rightarrow \lambda\varepsilon < \varepsilon_0) \\ &= f(x_0) + \widetilde{\varepsilon}\widetilde{\alpha}. \end{aligned}$$

Assim λh é uma direção de descida. Portanto $D(f; x_0)$ é um cone com vértice em 0. Verifiquemos agora que $D(f; x_0)$ é aberto, para isto considere $h \in D(f; x_0)$ e $\widetilde{h} \in U$. Tome α e ε_0 os mesmos da definição de direção de descida para h, então $f(x_0 + \varepsilon\bar{h}) \leq f(x_0) + \varepsilon\alpha$, para todo $\bar{h} \in U$ e todo $\varepsilon \in]0, \varepsilon_0[$. Portanto $\widetilde{h} \in D(f; x_0)$ e consequentemente $D(f; x_0)$ é aberto. □

Definição 3.3. *Seja Q uma **restrição de desigualdade**. Dizemos que $h \in E$ é uma **direção factível ou viável** para Q no ponto $x_0 \in Q$, se existem uma vizinhança U de h e $\varepsilon_0 > 0$ tais que para todo $\varepsilon \in]0, \varepsilon_0[$ e todo $\bar{h} \in U$ tem-se que,*

$$x_0 + \varepsilon\bar{h} \in Q.$$

Definição 3.4. *Uma **restrição de desigualdade** é dita **regular** no ponto x_0, se o conjunto das direções factíveis para Q em x_0 é convexo. Denotamos por $F(Q, x_0)$ ou K_F o conjunto de todas as direções factíveis.*

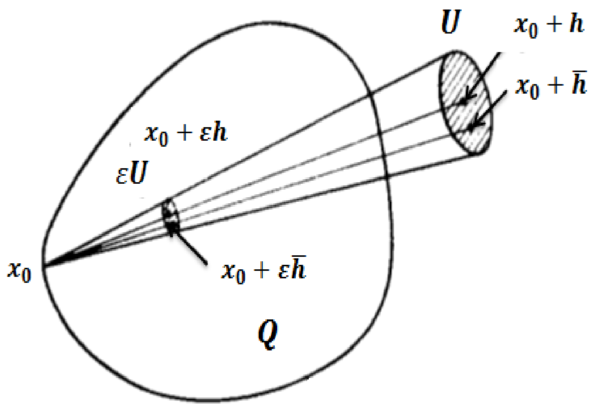


Figura 3.1: Direção de factível

Lema 3.2. *O conjunto de direções factíveis $F(Q; x_0)$ é um cone aberto com vértice em 0.*

Demonstração. Se $h \in F(Q; x_0)$, então $\lambda h \in F(Q; x_0)$ para $\lambda > 0$. Basta tomar V por

λU e $\widetilde{\varepsilon}_0 = \frac{\varepsilon_0}{\lambda}$. Logo

$$x_0 + \varepsilon\widetilde{h} = x_0 + \frac{\varepsilon}{\lambda}\lambda\bar{h} \in Q \ \forall\widetilde{h} \in V.$$

Assim $\lambda h \in F(Q; x_0)$. Portanto $F(Q; x_0)$ é um cone com vértice na origem.
Além disso, para qualquer $\widetilde{h} \in U$, tomando $\varepsilon \in]0, \varepsilon_0[$ e a mesma vizinhança U de modo que $x_0 + \varepsilon\bar{h} \in Q$ para toda direção $\bar{h} \in U$ segue que $\widetilde{h} \in F(Q; x_0)$, isto mostra que $F(Q; x_0)$ é aberto. □

Para restrições de igualdade, o conjunto de direções factíveis é vazio. Assim, torna-se necessário uma ligeira modificação em tal definição para este tipo de restrição.

Definição 3.5. *Seja $Q \subset E$ uma* ***restrição de igualdade****. Dizemos que um vetor $h \in E$ é* ***direção tangente*** *a Q no ponto $x_0 \in E$, se existem $\varepsilon_0 > 0$ e uma aplicação $r :]-\varepsilon_0, \varepsilon_0[\longrightarrow E$,tais que:*

a) $x(\varepsilon) = x_0 + \varepsilon h + r(\varepsilon) \in Q$;

b) $\frac{1}{\varepsilon}r(\varepsilon) \in V$ *para qualquer vizinhança V de 0 e para todo $\varepsilon > 0$ suficientemente pequeno, ou equivalentemente, $\|r(\varepsilon)\| = o(\varepsilon)$ para todo $\varepsilon > 0$ suficientemente pequeno.*

Denotamos por $T_{x_0}Q$ ou K_T ou $T(Q; x_0)$ ao conjuntos de todas as direções tangentes a Q no ponto x_0.

Definição 3.6. *Diremos que a restrição de igualdade Q é* ***regular*** *no ponto $x_0 \in Q$ se o conjunto $T(Q; x_0)$ é convexo.*

Lema 3.3. *O conjunto de todas as direções tangentes a Q no ponto x_0 geram um cone com vértice em na origem.*

Demonstração. Supondo $h \in T(Q; x_0)Q$ basta tomar $\tilde{\varepsilon}_0 = \lambda\varepsilon_0$ na definição de direção tangente e então para $\lambda > 0$, $\lambda h \in T(Q; x_0)$. □

Definição 3.7. *O vetor $h \in E$ é dito* ***direção tangente unilateral*** *a Q no ponto $x_0 \in Q$, se existem $\varepsilon_0 > 0$ e uma aplicação $r :]0, \varepsilon[\rightarrow E$ tais que:*

a) $x_0 + \varepsilon h + r(\varepsilon) \in Q$*;*

b) $\|r(\varepsilon)\| = o(\varepsilon)$*, para todo $\varepsilon > 0$ suficientemente pequeno.*

Denotamos por $T^+_{x_0}$ ou $T^+(Q; x_0)$ ao conjuntos de todas as direções tangentes unilaterais a Q no ponto x_0. O gráfico a seguir mostra uma direção tangente unilateral $h \in E$ em um ponto arbitrário $x_0 \in Q$:

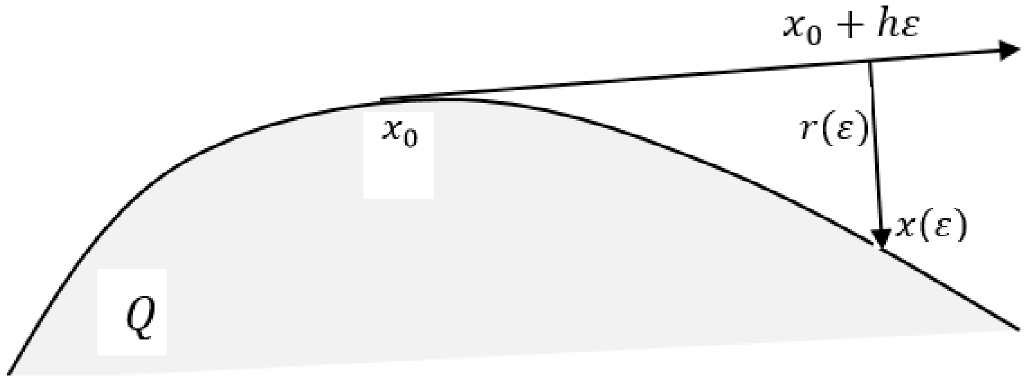


Figura 3.2: Direção tangente unilateral

Observação 3.1. *i) É claro que* $T_{x_0}Q \subset T^+_{x_0}Q$ *e* $T_{x_0}Q = T^+_{x_0}Q \cap T^-_{x_0}Q$*, onde* $T^-_{x_0}Q = -T^+_{x_0}Q$ *é definido como na Definição 3.7, mas com* $\varepsilon_0 < 0$ *e* $\varepsilon \to 0^-$.

ii) Se o conjunto $T_{x_0}Q$ *é um subespaço vetorial de* E*, então ele se chamará* ***espaço tangente*** *a* Q *no ponto* x_0.

iii) Toda direção factível é também uma direção tangente, porém a recíproca é falsa. De fato, se $h \in F(Q; x_0)$*, então existe* $\varepsilon_0 > 0$ *tal que* $x_0 + \varepsilon\bar{h} \in Q$*, com* $\bar{h} \in V$*, onde* V *é uma vizinhança de* h *e* $\varepsilon \in]0, \varepsilon_0[$*. Em particular* $x_0 + \varepsilon h \in Q$*, para todo* $\varepsilon \in]-\varepsilon_0, \varepsilon_0[$*. Isto implica que* h *é uma direção tangente a* Q *no ponto* x_0 *com* $r(\varepsilon) = 0$.

Como contra exemplo, considere $E = \mathbb{R}^2$*,* $Q = \{(x, y) \in \mathbb{R}^2 : y \geq x^2\}$ *e* $x_0 = (0, 0)$*. Logo* $h = (1, 0) \in T(Q; x_0)$*, mas* $h \notin F(Q; x_0)$.

Exemplo 3.1. *Seja* $E = \mathbb{R}^2$*,* $Q = \{(x_1, x_2) : x_1 \geq 0\}$*. Então* $T^+_{(0,0)}Q = Q$*;* $T_{(0,0)}Q = \{(0, b) : b \in \mathbb{R}\}$*;* $T_{(1,0)}Q = T^+_{(1,0)}Q = \mathbb{R}^2$.
De fato, tomando $x_0 = (0, 0)$*,* $h = (h_1, h_2)$ *e* $r(\varepsilon) = (r_1(\varepsilon), r_2(\varepsilon))$ *determinaremos quais os vetores* h *que satisfazem as condições da definição de vetor tangente:*
$x_0 + \varepsilon h + r(\varepsilon) \in Q \Longleftrightarrow (0, 0) + \varepsilon(h_1, h_2) + (r_1(\varepsilon), r_2(\varepsilon)) = (\varepsilon h_1 + r_1(\varepsilon).\varepsilon h_2 + r_2(\varepsilon)) \in Q$.
Daí $\varepsilon h_1 + r_1(\varepsilon) \geq 0$,

i) Se $\varepsilon \to 0^+$*,* $h_1 \geq -\frac{r_1(\varepsilon)}{\varepsilon}$*. Assim,* $h_1 \geq 0$*. Portanto* $T^+_{(0,0)}Q = \{(x_1, x_2) : x_1 \geq 0\}$.

ii) Se $\varepsilon \to 0^-$*,* $h_1 \leq -\frac{r_1(\varepsilon)}{\varepsilon}$*. Então* $h_1 \leq 0$ *e* $-T^-_{(0,0)}Q = \{(x_1, x_2) : x_1 \leq 0\}$.

Portanto, $T_{(0,0)}Q = T^+_{(0,0)}Q \cap T^-_{(0,0)}Q = \{(0, b) : b \in \mathbb{R}\}$.
Considerando $x_0 = (1, 0)$*, teremos:*

$$(1, 0) + \varepsilon(h_1, h_2) + (r_1(\varepsilon), r_2(\varepsilon)) \in Q \Longleftrightarrow \varepsilon h_1 + r_1(\varepsilon) + 1 \geq 0 \Rightarrow \varepsilon h_1 + r_1(\varepsilon) \geq -1.$$

Por conseguinte,

$$\begin{cases} h_1 & \geq & \frac{-1-r_1(\varepsilon)}{\varepsilon}, \ \textit{quando } \varepsilon > 0, \ \textit{donde segue que } h_1 \geq -\infty, \ \textit{se } \varepsilon \to 0^+; \\ h_1 & \leq & \frac{-1-r_1(\varepsilon)}{\varepsilon}, \ \textit{quando } \varepsilon < 0, \ \textit{donde segue que } h_1 \leq +\infty, \ \textit{se } \varepsilon \to 0^-. \end{cases}$$

Consequentemente, $T_{(1,0)}Q = T^+_{(1,0)}Q = \mathbb{R}^2$.

O esboço a seguir mostra os cones tangentes $T^+_{(0,0)}Q$, $T_{(0,0)}Q$, $T_{(1,0)}Q$ e $T^+_{(1,0)}Q$.

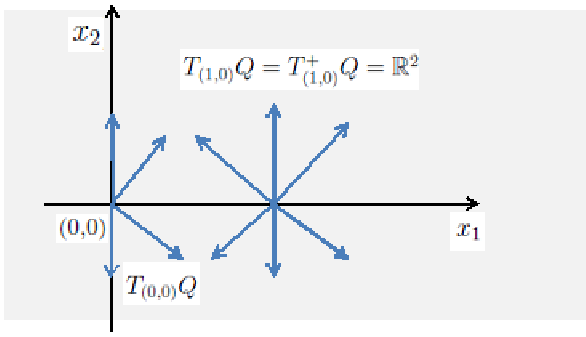


Figura 3.3: Direção tangente unilateral

O resultado a seguir, é de fundamental importância na obtenção da condição necessária de otimalidade para o problema (P_0), conhecida como equação de Euler-Lagrange.

Lema 3.4. *(Dubovitskii-Milyutin). Sejam* $K_1, ..., K_n, K_{n+1}$ *cones convexos com vértice em* 0*, onde* $K_1, ..., K_n$ *são abertos. Então* $\bigcap_{i=1}^{n+1} K_i = \emptyset$ *se, e somente se, existem funcionais lineares* $f_i \in K_i^*$, $i = 1, ..., n+1$, *não todos nulos, tais que*

$$f_1 + f_2 + ... + f_n + f_{n+1} = 0.$$

Demonstração. ***Necessária.*** Seja $\bigcap\limits_{i=1}^{n} K_i = K \neq \emptyset$, então $K_{n+1} \bigcap K = \emptyset$. Como K é aberto e com vértice em 0, aplicando o teorema de separação, Teorema 2.6, e Teorema 2.1 temos que existe $f \in E'$, $f \neq 0$, tal que

$$f(x) \geq 0 \ \forall x \in K,$$

e

$$f(x) \leq 0 \ \forall x \in K_{n+1}.$$

Como $K_1, \ldots, K_n$ são cones abertos e convexos, segue do Teorema (2.2) que

$$K^* = \left(\bigcap_{i=1}^{n} K_i\right)^* = \sum_{i=1}^{n} K_i^*,$$

i.e., se $f \in K^*$ então $f = \sum_{i=1}^{n} f_i$, onde $f_i \in K_i^*$ para $i = 1, 2, \ldots, n$.

Defina $f_{n+1} = -f$. Então $f_{n+1} \in K_{n+1}^*$, $f_{n+1} \neq 0$ e $f_1 + f_2 + \cdots + f_n + f_{n+1} = 0$. Mas se $\bigcap_{i=1}^{n} K_i = \emptyset$, existe $s \in \{1, 2, \ldots, n-1\}$ tal que $\bigcap_{i=1}^{s} K_i \neq \emptyset$ e $\bigcap_{i=1}^{s+1} K_i = \emptyset$. Aplicando o resultado anterior trocando n por s, obtemos $f_i \in K_i^*; i = 1, 2, \ldots, s+1$, $f_{s+1} \neq 0$ tal que $f_1 + f_2 + \cdots + f_n + f_{s+1} = 0$. Tomando $f_{s+2} = \cdots = f_{n+1} = 0$ obtemos os $n+1$ funcionais desejados $f_1, \ldots, f_{n+1}$.

Suficiência. Seja $f_1 + f_2 + \cdots + f_n + f_{n+1} = 0$, $f_i \in K^*$, nem todos nulos. Suponhamos pelo absurdo que, $\bigcap_{i=1}^{n+1} K_i \neq \emptyset$, então existe $x_0 \in K_i$, $i = 1, \ldots, n+1$. Ao mesmo tempo deve existir $f_j \neq 0$, para algum $1 \leq j \leq n$ (caso contrário teríamos $f_{n+1} = -\sum_{i=1}^{n} f_i = 0$) e então $f_j(x_0) > 0$ (desde que K_i é aberto), mas $0 = (f_1 + f_2 + \cdots + f_{n+1})(x_0) \geq f_j(x_0) > 0$ o que é um absurdo. Portanto, $\bigcap_{i=1}^{n+1} K_i = \emptyset$. $\square$

Observação 3.2. *No decorrer da demostração do Lema 3.4 afirmamos que se K_i é aberto para todo $i = 1, \ldots, n$, então $f_j(x_0) > 0$. De fato, note que, se $x \in K_j$, implica que $f_j(x) \geq 0$. Portanto $K_j \subset \{x : f_j(x) \geq 0\}$. K_j é aberto se, $K_j \subset int\{x \in E : f_j(x) \geq 0\} = \{x \in E : f_j(x) > 0\}$.*
Logo, dado $x \in K_j$ existe um aberto U tal que $x \in U \subset K_j$, mas $K_j \subset \{x : f_j(x) \geq 0\}$, então $x \in X \subset \{x : f_j(x) \geq 0\}$. Logo $x \in int\{x : f_j(x) \geq 0\}$.

Teorema 3.1. *(Dubovitskii-Milyutin). Seja $f : U \subset E \to \mathbb{R}$ um funcional, numa vizinhança U de x_0, e considere o Problema (P_0). Se,*

(i) $f(x)$ é regular de descida em x_0 com direções de descida no cone K_0;

(ii) As restrições de desigualdades em Q_i, $i = 1, ..., n$, são regulares em x_0, com direções factíveis no cone K_i;

(iii) A restrição de igualdade em Q_{n+1}, é também regular em x_0 com direções tangentes no cone K_{n+1}.

Então existem funcionais lineares contínuos $f_i \in K_i^*$, $i = 0, \ldots, n+1$, nem todos nulos, tais que satisfazem a equação de Euler-Lagrange

$$f_0 + f_1 + \cdots + f_n + f_{n+1} = 0.$$

Demonstração. Primeiramente vamos provar que a condição necessária para o funcional ter um mínimo em x_0 é que,

$$\bigcap_{i=0}^{n+1} K_i = \emptyset,$$

ou seja, nenhuma direção de descida do funcional $f(x)$ pode ser factível para todas as restrições. Então suponhamos que isto é falso, assim existe $h \in K_i$, $i = 0, 1, \ldots, n+1$. Pela definição de K_i, $i = 0, 1, \ldots, n$ existem, uma vizinhança U de h e $\varepsilon_0 > 0$ tal que para qualquer $0 < \epsilon < \varepsilon_0$ e $\bar{h} \in U$, o vetor $x_0 + \varepsilon\bar{h} \in \bigcap_{i=1}^{n} Q_i$ e satisfaz a desigualdade

$$f(x_0 + \varepsilon\bar{h}) \leq f(x_0) + \varepsilon\alpha,$$

para algum $\alpha < 0$. Consideremos agora o vetor $x(\varepsilon) = x_0 + \varepsilon h + r(\varepsilon) \in Q_{n+1}$ como na definição 3.5 de direção tangente, e seja $\varepsilon_1 > 0$ tal que $\frac{1}{\varepsilon}r(\varepsilon) \in U \setminus h$ com $0 < \varepsilon \leq \varepsilon_1$, ou seja, $\bar{h}(\varepsilon) = h + \frac{1}{\varepsilon}r(\varepsilon) \in U$, para $0 < \varepsilon \leq \varepsilon_1$. Então quando $0 < \varepsilon < \min\{\varepsilon_0, \varepsilon_1\}$ os vetores $x(\varepsilon) = x_0 + \varepsilon\bar{h}(\varepsilon) = x_0 + \varepsilon h + r(\varepsilon)$ estão ao mesmo tempo em $\bigcap_{i=0}^{n} Q_i$ e Q_{n+1}. Assim eles satisfazem as restrições, mas eles também satisfazem a desigualdade

$$f(x(\epsilon)) = f(x_0 + \epsilon\bar{h}(\epsilon)) \leq f(x_0) + \epsilon\alpha < f(x_0),$$

o que contradiz o fato de x_0 ser um mínimo local. Portanto

$$\bigcap_{i=0}^{n+1} K_i = \emptyset.$$

Como cada K_i é um cone aberto e convexo para todo $i = 0, \ldots, n$, e K_{n+1} é convexo, então pelo Lema 3.4 existem funcionais lineares contínuos $f_i \in E'$ para $i = 0, \ldots, n+1$, não todos nulos, tais que

$$f_0 + f_1 + \cdots + f_n + f_{n+1} = 0.$$

□

Observação 3.3. *Uma condição suficiente para garantir que* $f_0 \neq 0$ *é que exista pelo menos uma direção factível e uma direção tangente, i.e.,*

$$\bigcap_{i=1}^{n+1} Q_i \neq \emptyset.$$

A conclusão é válida para qualquer f_i, $i = 1, \ldots, n, n+1$.

O teorema anterior proporciona condições necessárias de otimalidade para um problema extremal, as quais são uma generalização da Regra dos Multiplicadores de Lagrange da análise clássica, da equação de Euler-Lagrange no Cálculo de Variações e do Princípio de Máximo de Pontryagin para Problemas de Controle Ótimo, que será motivo de discussão no próximo capítulo. Note que para obter essas condições necessárias, para um problema específico, é necessário determinar as direções de descida, direções fatíveis e direções tangentes e seus respetivos duais.

3.2 Cálculo de Cones e seus Duais

Nesta seção faremos uma breve seleção de resultados, os quais nos permitem exibir os cones de direções de descida, direções factíveis e direções tangentes associados ao problema que estudaremos no próximo capítulo.

3.2.1 Direções de Descida

Como veremos a seguir, as derivadas direcionais fornecem um tópico conveniente para a construção do cone de direções de decrescimento (descida).

Definição 3.8. *Um funcional $f(x)$ num espaço linear normado E diz-se que possui derivada $f'(x_0, h)$ no ponto x_0 na direção h se,*

$$\lim_{\epsilon \to +0} \frac{f(x_0 + \epsilon h) - f(x_0)}{\epsilon} = f'(x_0, h)$$

existe.
Esta derivada é conhecido na bibliografia como a derivada de Gatêux.

Observação 3.4. *Se $E = \mathbb{R}$ (f é uma função de variável real), então a existência de $f'(x_0, h)$ é equivalente à existência das derivadas laterais à direita ($h > 0$) e à esquerda ($h < 0$) de f no ponto x_0.*

Teorema 3.2. *Se $h \in D(f; x_0)$ e $f'(x_0, h)$ existe, então $f'(x_0, h) < 0$.*

Demonstração. Se $h \in D(f; x_0)$, segue da definição de direção de descida que, $f(x_0+\varepsilon h) \leq f(x_0) + \varepsilon\alpha$ para ε suficientemente pequeno, logo

$$f'(x_0, h) = \lim_{\varepsilon \to +0} \frac{f(x_0 + \varepsilon h) - f(x_0)}{\varepsilon} \leq \alpha < 0.$$

□

Agora consideremos uma outra classe de funcionais com a propriedade de que são regularmente decrescentes (de descida) em um ponto, e que o cone de direções de descida são fáceis de determinar. Vamos definir então a noção de derivada para operadores definidas sob espaços de Banach.

Definição 3.9. *Sejam E_1 e E_2 espaços de Banach e $F : E_1 \to E_2$ um operador (geralmente não linear). F é dito* ***Fréchet diferenciável*** *em $x_0 \in E_1$ se existe um operador linear $\Lambda : E_1 \to E_2$ tal que para todo $h \in E_1$*

$$F(x_0 + h) = F(x_0) + \Lambda h + r(x_0, h), \tag{3.1}$$

onde

$$\|r(x_0, h)\| = o(\|h\|). \tag{3.2}$$

O operador linear Λ é denotado por $F'(x_0)$ ou $DF(x_0)$. Quando $E_2 = \mathbb{R}$, Λ é um funcional linear.

Teorema 3.3. *Se $f(x)$ é diferenciável em $x_0 \in E$ no sentido de Frechet, então $D(f; x_0) = \{h \in E : f'(x_0)h < 0\}$ e $f(x)$ é regular de descida em x_0.*

Demonstração. Como $f(x)$ é diferenciável em x_0, então $f(x)$ possui derivada direcional em qualquer direção e do Teorema 3.2 segue que $D(f; x_0) \subset \{h \in E : f'(x)h < 0\}$. Por outro lado, da diferenciabilidade de $f(x)$ em x_0, temos

$$f(x_0 + \eta) = f(x_0) + f'(x_0)\eta + o(\|\eta\|)$$

Tomando $\eta = \varepsilon \bar{h}$, para $\varepsilon > 0$ suficientemente pequeno e $\bar{h} \in E$ pertence a uma vizinhança de h, e tendo em conta que $f'(x_0)h < 0$, segue que h é uma direção de descida, isto é, $h \in D(f; x_0)$. Assim mostramos a igualdade desejada.

Agora, como $\{h \in E : f'(x_0)h < 0\}$ é convexo, pois é um semi-espaço o conjunto não vazio, então $D(f; x_0)$ é regularmente de descida. □

Exemplo 3.2. *Consideremos $E = C([0, T]; \mathbb{R}^n)$, $x = x(t) \in E$ e o funcional integral*

$$F(x) = \int_0^T \Phi(x(t), t)dt$$

onde $\Phi(x, t)$ é contínua em (x, t), diferenciável com respeito a x e $\Phi_x(x, t)$ contínua em (x, t). Então $F(x)$ é diferenciável e

$$(F'(x_0), h) = \int_0^T \langle \Phi_x(x_0(t), t), h(t) \rangle \, dt,$$

$$D(F, x_0) = \left\{ h \in E : \int_0^T \langle \Phi_x(x_0(t), t), h(t) \rangle \, dt < 0 \right\}.$$

De fato, como $\Phi(\cdot, t)$ é diferenciável com respeito a x pode-se usar o Teorema de Valor Médio,

$$\begin{aligned} F(x_0+h) - F(x_0) &= \int_0^T [\Phi(x_0(t)+h(t),t) - \Phi(x_0(t),t)]\, dt \\ &= \int_0^T \langle \Phi_x(x_0(t)+\theta(t)h(t),t), h(t)\rangle\, dt + \int_0^T \langle \Phi_x(x_0(t),t), h(t)\rangle\, dt \\ &\quad - \int_0^T \langle \Phi_x(x_0(t),t), h(t)\rangle\, dt \\ &= \int_0^T \langle \Phi_x(x_0(t),t), h(t)\rangle\, dt + \\ &\quad \int_0^T \langle \Phi_x\left(x_0(t)+\theta(t)h(t),t\right) - \Phi_x(x_0(t),t), h(t)\rangle\, dt \end{aligned}$$

onde, $\theta(t) \in]0,1[$. Como $\Phi_x(x(t),t)$ é contínua em todo (x,t), então é uniformemente contínua em $M \times [0,T]$, onde

$$M = \{x \in \mathbb{R}^n : |x| \le \|x_0\| + \|h\|\}$$

é compacto. Logo $\forall \varepsilon > 0$, $\exists\, \delta > 0$ tal que

$$|\Phi_x(x_1,t) - \Phi_x(x_2,t)| \le \varepsilon,$$

para todo $x_1, x_2 \in M$, $|x_1 - x_2| \le \delta$ e $t \in [0,T]$.
Claro que $x_0(t), x_0 + \theta(t)h(t) \in M$, portanto $|\Phi_x(x_0(t)+\theta(t)h(t),t) - \Phi_x(x_0(t),t)| \le \varepsilon$ para $|h(t)| \le \delta$, i.e., para $\|h\| \le \delta$. Por conseguinte

$$\max_{0\le t\le T} |\Phi_x(x_0(t)+\theta(t)h(t),t) - \Phi_x(x_0(t),t)| \to 0 \tag{3.3}$$

quando $\|h\| \to 0$, ou seja,

$$\left| \int_0^T \langle \Phi_x(x_0(t)+\theta(t)h(t),t) - \Phi_x(x_0(t),t), h(t)\rangle\, dt \right|$$

$$\begin{aligned} &\le \int_0^T |\langle \Phi_x(x_0(t)+\theta(t)h(t),t) - \Phi_x(x_0(t),t), h(t)\rangle|\, dt \\ &\le \int_0^T |\Phi_x(x_0(t)+\theta(t)h(t),t) - \Phi_x(x_0(t),t)|\, |h(t)| dt \\ &\le \int_0^T \max_{0\le t\le T} |\Phi_x(x_0(t)+\theta(t)h(t),t) - \Phi_x(x_0(t),t)|\, \|h\| dt \\ &= T \max_{0\le t\le T} |\Phi_x(x_0(t)+\theta(t)h(t),t) - \Phi_x(x_0(t),t)| \|h\| \\ &= o(\|h\|). \end{aligned}$$

Portanto,

$$F(x_0+h)-F(x_0)=(F'(x_0),h)+o(\|h\|).$$

Assim $F(x)$ é diferenciável em x_0, onde $(F'(x_0),h)=\int_0^T\langle\Phi_x(x_0(t),t),h(t)\rangle$. Logo pelo Teorema 3.3, $F(x)$ é regularmente de descida e

$$D(F,x_0)=\{h:(F'(x_0),h)<0\}=\left\{h:\int_0^T(\Phi_x(x_0(t),t),h(t))dt<0\right\}.$$

Exemplo 3.3. *Sejam* $E=C([0,T];\mathbb{R}^n)\times L_\infty([0,T];\mathbb{R}^n)$, $(x,u)=(x(t),u(t))\in E$ *e o funcional integral*

$$F(x,u)=\int_0^T\Phi(x(t),u(t),t)dt.$$

Suponha que $\Phi(x,u,t)$ *é contínua em* (x,u,t), *diferenciável com respeito a* x,u *e* $\Phi_x(x,u,t)$, $\Phi_u(x,u,t)$ *são contínuas em* (x,u,t). *Então* $F(x,u)$ *é diferenciável e*

$$F'(x,u)(\bar{x},\bar{u})=\int_0^T\left[\langle\Phi_x(x(t),u(t),t),\bar{x}(t)\rangle+\langle\Phi_u(x(t),u(t),t),\bar{u}(t)\rangle\right]dt,$$

$$D\left(F,(x,u)\right)=\left\{(\bar{x},\bar{u})\in E:F'(x,u)(\bar{x},\bar{u})<0\right\}.$$

De fato, como $\Phi(x,u,t)$ é continuamente diferenciável em x,u, usando o Teorema de Valor Médio para $\Phi(\cdot,\cdot,t)$ com $t\in[0,T]$ fixo, existem $\theta_1(t),\ \theta_2(t)\in]0,1[$ tais que

$$\Phi(x+\bar{x},u+\bar{u},t)-\Phi(x,u,t)=\langle\Phi_x\left(x+\theta_1(t)\bar{x},u,t\right),\bar{x}(t)\rangle+\langle\Phi_u\left(x,u+\theta_2(t)\bar{u},t\right),\bar{u}(t)\rangle.\tag{3.4}$$

Então

$F(x+\bar{x},u+\bar{u})-F(x,u)$

$$\begin{aligned}
&= \int_0^T[\Phi(x+\bar{x},u+\bar{u},t)-\Phi(x,u,t)]dt\\
&\overset{(3.4)}{=} \int_0^T\left[\langle\Phi_x\left(x+\theta_1(t)\bar{x},u,t\right),\bar{x}(t)\rangle+\langle\Phi_u\left(x,u+\theta_2(t)\bar{u},t\right),\bar{u}(t)\rangle\right]dt\\
&\quad+\int_0^T\left[\langle\Phi_x(x(t),u(t),t),\bar{x}(t)\rangle+\langle\Phi_u(x(t),u(t),t),\bar{u}(t)\rangle\right]dt\\
&\quad-\int_0^T\left[\langle\Phi_x(x(t),u(t),t),\bar{x}(t)\rangle+\langle\Phi_u(x(t),u(t),t),\bar{u}(t)\rangle\right]dt\\
&= \int_0^T\left[\langle\Phi_x(x(t),u(t),t),\bar{x}(t)\rangle+\langle\Phi_u(x(t),u(t),t),\bar{u}(t)\rangle\right]dt\\
&\quad+\int_0^T\langle\Phi_x\left(x(t)+\theta_1(t)\bar{x}(t),u(t),t\right)-\Phi_x(x(t),u(t),t),\bar{x}(t)\rangle\,dt\\
&\quad+\int_0^T\langle\Phi_u\left(x(t),u(t)+\theta_2(t)\bar{u}(t),t\right)-\Phi_u(x(t),u(t),t),\bar{u}(t)\rangle\,dt.
\end{aligned}$$

É fácil ver que,

$$\int_0^T \langle \Phi_x\left(x(t)+\theta_1(t)\bar{x}(t),u(t),t\right)-\Phi_x(x(t),u(t),t),\bar{x}(t)\rangle\, dt+$$

$$\int_0^T \langle \Phi_u\left(x(t),u(t)+\theta_2(t)\bar{u}(t),t\right)-\Phi_u(x(t),u(t),t),\bar{u}(t)\rangle\, dt = o(\|(\bar{x},\bar{u})\|).$$

Logo $F(x,u)$ é diferenciável, obtendo assim as conclusões do exemplo. De fato,

$$|\int_0^T \langle \Phi_x\left(x(t)+\theta_1(t)\bar{x}(t),u(t),t\right)-\Phi_x(x(t),u(t),t),\bar{x}(t)\rangle dt+$$

$$\begin{aligned}
& \int_0^T \langle \Phi_u\left(x(t),u(t)+\theta_2(t)\bar{u}(t),t\right)-\Phi_u(x(t),u(t),t),\bar{u}(t)\rangle\, dt| \\
\leq\ & \int_0^T \left|\Phi_x\left(x(t)+\theta_1(t)\bar{x}(t),u(t),t\right)-\Phi_x(x(t),u(t),t)\right| \|\bar{x}\|_C\, dt \\
& + \int_0^T \left|\Phi_u\left(x(t),u(t)+\theta_2(t)\bar{u},t\right)-\Phi_u(x(t),u(t),t)\right| \|\bar{u}\|_\infty\, dt \\
\leq\ & \|\bar{x}\|_C \int_0^T \max_{t\in[0,T]} \left|\Phi_x\left(x(t)+\theta_1(t)\bar{x}(t),u(t),t\right)-\Phi_x(x(t),u(t),t)\right| dt \\
& + \|\bar{u}\|_\infty \int_0^T \max_{t\in[0,T]} \left|\Phi_u\left(x(t),u(t)+\theta_2(t)\bar{u},t\right)-\Phi_u(x(t),u(t),t)\right| dt.
\end{aligned}$$

Esta última expressão converge a 0, quando $\|\bar{x}\| \to 0$ e $\|\bar{u}\| \to 0$ (do modo similar à obtenção de (3.3) no exemplo anterior). Portanto fica provado a afirmação.

Teorema 3.4. *Seja $f(x)$ um funcional convexo e contínuo em um espaço de Banach E, então:*

i) *$f(x)$ tem derivada direcional em qualquer ponto, e em qualquer direção;*

ii) $f(x_0+h) \geq f(x_0)+f'(x_0,h)$;

iii) $D(f;x_0)=\{h\in E : f'(x_0,h)<0\}$;

iv) *$f(x)$ é regular de descida em qualquer ponto $x_0\in E$.*

Demonstração. i) Sejam $x_0\in E$ e $h\in E$ fixo, e considere a função real

$$\varphi(\varepsilon)=f(x_0+\varepsilon h).$$

Como $f(x)$ é convexa, segue que $\varphi(\varepsilon)$ é também convexa, De fato, para qualquer $0 < \lambda < 1$ e $\varepsilon_1, \varepsilon_2 \in \mathbb{R}$ temos o seguinte:

$$\begin{aligned}\varphi\left(\lambda\varepsilon_1 + (1-\lambda)\varepsilon_2\right) &= f(x_0 + (\lambda\varepsilon_1 + (1-\lambda)\varepsilon_2)h) \\ &= f\left(\lambda(x_0 + \varepsilon_1 h) + (1-\lambda)(x_0 + \varepsilon_2 h)\right) \\ &\leq \lambda f(x_0 + \varepsilon_1 h) + (1-\lambda) f(x_0 + \varepsilon_2 h) \\ &= \lambda\varphi(\varepsilon_1) + (1-\lambda)\varphi(\varepsilon_2).\end{aligned}$$

Donde, para qualquer $0 < \lambda < 1$ temos

$$\varphi(\lambda\varepsilon) \leq \lambda\varphi(\varepsilon) + (1-\lambda)\varphi(0).$$

Daí, para $\varepsilon > 0$.

$$\frac{\varphi(\lambda\varepsilon) - \varphi(0)}{\lambda\varepsilon} \leq \frac{\varphi(\varepsilon) - \varphi(0)}{\varepsilon},$$

portanto, $\dfrac{\varphi(\varepsilon) - \varphi(0)}{\varepsilon}$ é monótona decrescente para $\varepsilon \to +0$. Além disso,

$$\begin{aligned}\varphi(0) = \varphi\left(\frac{-\varepsilon}{1+\varepsilon} + \frac{\varepsilon}{1+\varepsilon}\right) &\leq \frac{\varepsilon}{1+\varepsilon}\varphi(-1) + \frac{1}{1+\varepsilon}\varphi(\varepsilon) \\ &\Rightarrow (1+\varepsilon)\varphi(0) \leq \varepsilon\varphi(-1) + \varepsilon\varphi(\varepsilon) \\ &\Rightarrow \varphi(0) - \varphi(-1) \leq \frac{\varphi(\varepsilon) - \varphi(0)}{\varepsilon}.\end{aligned}$$

Assim, $\dfrac{\varphi(\varepsilon) - \varphi(0)}{\varepsilon}$ é limitada inferiormente para $\varepsilon > 0$.
Portanto, $\lim_{\varepsilon \to 0} \dfrac{\varphi(\varepsilon) - \varphi(0)}{\varepsilon}$ existe. Logo, segue da definição de derivada direcional que,

$$f'(x_0, h) = \lim_{\varepsilon \to +0} \frac{\varphi(\varepsilon) - \varphi(0)}{\varepsilon}.$$

Por conseguinte, a derivada direcional existe em qualquer ponto e em qualquer direção.

ii) Do item i) temos que

$$\frac{\varphi(\varepsilon) - \varphi(0)}{\varepsilon}$$

é monótona decrescente, isto que implica

$$\varphi(1) - \varphi(0) \geq \lim_{\varepsilon \to 0^+} \frac{\varphi(\varepsilon) - \varphi(0)}{\varepsilon},$$

ou seja,

$$f(x_0 + h) \geq f(x_0) + f'(x_0, h).$$

iii) Pelo Teorema 3.2 temos que $D(f; x_0) \subset \{h \in E : f'(x_0, h) < 0\}$. Resta provar a inclusão contrária. Considere $h \in E$ tal que $f'(x_0, h) < 0$. Segue da definição de derivada direcional que existe $\varepsilon_0 > 0$ tal que

$$f(x_0 + \varepsilon_0 h) < f(x_0).$$

Tome

$$\delta = f(x_0) - f(x_0 + \varepsilon_0 h) > 0. \tag{3.5}$$

Como $f(x)$ é contínuo, considere U uma vizinhança do vetor h tal que para todo $\bar{h} \in U$ se verifica

$$|f(x_0 + \varepsilon_0 \bar{h}) - f(x_0 + \varepsilon_0 h)| \leq \frac{\delta}{2}.$$

Então,

$$f(x_0 + \varepsilon_0 \bar{h}) \leq f(x_0 + \varepsilon_0 h) + \frac{\delta}{2},$$

e por (3.5),

$$f(x_0 + \varepsilon_0 \bar{h}) \leq f(x_0) - \delta + \frac{\delta}{2} = f(x_0) - \frac{\delta}{2}. \tag{3.6}$$

Agora da condição de convexidade, para $0 < \varepsilon < \varepsilon_0$ temos que

$$\begin{aligned}
f(x_0 + \varepsilon\bar{h}) &= f(x_0 - \frac{\varepsilon}{\varepsilon_0}x_0 + \frac{\varepsilon}{\varepsilon_0}x_0 + \frac{\varepsilon}{\varepsilon_0}\varepsilon_0\bar{h}) \\
&= f\left(\left(1 - \frac{\varepsilon}{\varepsilon_0}\right)x_0 + \frac{\varepsilon}{\varepsilon_0}\left(x_0 + \varepsilon_0\bar{h}\right)\right) \\
&\leq \left(1 - \frac{\varepsilon}{\varepsilon_0}\right)f(x_0) + \frac{\varepsilon}{\varepsilon_0}f(x_0 + \varepsilon_0\bar{h}),
\end{aligned}$$

e por (3.6),

$$\begin{aligned}
f(x_0 + \varepsilon\bar{h}) &\leq f(x_0) - \frac{\varepsilon}{\varepsilon_0}f(x_0) + \frac{\varepsilon}{\varepsilon_0}\left(f(x_0) - \frac{\delta}{2}\right) \\
&= f(x_0) - \frac{\varepsilon\delta}{2\varepsilon_0}.
\end{aligned}$$

Ou seja, a condição de direção de descida é satisfeita para $\alpha = -\frac{\delta}{2\varepsilon_0}$, portanto $h \in D$.

vi) Para mostrar que $f(x)$ é regular de descida vejamos que $D(f; x_0)$ é convexo. Sejam h_1, $h_2 \in D(f; x_0)$ e $\lambda \in [0, 1]$. Como $f(x)$ é convexo temos que

$$\begin{aligned}
\frac{f(x_0 + \varepsilon h) - f(x_0)}{\varepsilon} &= \frac{f\left(x_0 + \varepsilon(\lambda h_1 + (1-\lambda)h_2)\right) - f(x_0)}{\varepsilon} \\
&\leq \frac{\lambda f(x_0 + \varepsilon h_1) + (1-\lambda)f(x_0 + \varepsilon h_2) - f(x_0)}{\varepsilon} \\
&= \frac{\lambda[f(x_0 + \varepsilon h_1) - f(x_0)] + (1-\lambda)[f(x_0 + \varepsilon h_2) - f(x_0)]}{\varepsilon},
\end{aligned}$$

tomando limite quando $\varepsilon \to \infty$ esta ultima desigualdade temos que,

$$f'(x_0, h) \leq \lambda f'(x_0, h_1) + (1 - \lambda) f'(x_0, h_2).$$

Como h_1, $h_2 \in D(f; x_0)$ então $f'(x_0, h_1) < 0$ e $f'(x_0, h_2) < 0$. Assim, $f'(x_0, h) < 0$, ou seja, $h = \lambda h_1 + (1 - \lambda) h_2 \in D(f; x_0)$. Portanto $D(f; x_0)$ é convexo, e consequentemente $f(x)$ regular de descida. □

3.2.2 Direções Factíveis

Nas seções anteriores denotamos por $F(Q; x_0)$ ao cone de direções factíveis de um conjunto Q no ponto x_0 e por $D(f; x_0)$ ao cone direções de descida do funcional $f(x)$ no ponto x_0 e vamos considerar o caso em que o conjunto Q é definido por algum funcional $f(x)$, isto é,

$$Q = \{x \in E : f(x) \leq f(x_0)\}.$$

Observação 3.5.

i) *Se $x_0 \in int(Q)$, então $F(Q; x_0) = E$, assim os pontos interessantes são os pontos da fronteira de Q, ou seja $x \in Fr(Q)$.*

ii) *O conjunto Q deve ter pontos interiores, pois caso contrario é trivial, $F(Q; x_0) = \emptyset$.*

iii) *Se $f(x)$ é contínuo não faz sentido estudar o caso mais geral $Q = \{x : f(x) \leq \lambda)\}$, com $\lambda \neq f(x_0)$, pois se $f(x_0) < \lambda$ então $x_0 \in int(Q)$ e se $f(x_0) > \lambda$ então $x_0 \notin Q$, ou seja, x_0 não é um ponto factível.*

Lema 3.5. *O cone das direções de descida está contido no cone das direções factíveis, isto é, $D(f; x_0) \subset F(Q; x_0)$.*

Demonstração. Se $h \in D(f; x_0)$ então $f(x_0 + \varepsilon \bar{h}) \leq f(x_0) + \varepsilon \alpha < f(x_0)$ $\forall \bar{h} \in U$, $\alpha < 0$ e $\forall\ \varepsilon \in]0, \varepsilon_0[$. Portanto, $x_0 + \varepsilon \bar{h} \in Q$ $\forall \bar{h}$ e $\forall \varepsilon \in]0, \varepsilon_0[$, donde $h \in F(Q; x_0)$. □

A pergunta natural que surge é, quando $F(Q; x_0) = D(f; x_0)$?. O teorema a seguir responde a questão.

Teorema 3.5. *Seja $f(x)$ diferenciável em x_0 em qualquer direção e $f'(x_0, h)$ convexo em h. Se existe $\tilde{h}$ tal que $f'(x_0, \tilde{h}) < 0$, então $F(Q; x_0) \subset \{h : f'(x_0, h) < 0\} = D(f; x_0)$.*

Demonstração. Seja $h \in F(Q; x_0)$. Então existe $\varepsilon_0 > 0$ tal que $x_0 + \varepsilon h \in Q$ $\forall \varepsilon \in]0, \varepsilon_0)[$, mas $Q = \{x : f(x) \leq f(x_0)\}$, daí segue que

$$f(x_0 + \varepsilon h) \leq f(x_0) \Longrightarrow \frac{f(x_0 - \varepsilon h) - f(x_0)}{\varepsilon} \leq 0.$$

Logo, $f'(x_0, h) \leq 0$.
Além disso, o cone $F(Q; x_0)$ é aberto (Lema 3.2), logo existe uma vizinhança U de h tal que $U \subseteq F(Q; x_0)$. Tome $\gamma > 0$ tal que, $h_\gamma = h + \gamma(h - \tilde{h}) \in U$, ou seja $h = \frac{1}{\gamma+1} h_\gamma + \frac{\gamma}{\gamma+1}\tilde{h}$. Usando a convexidade de $f'(x_0, h)$ temos o seguinte:

$$f'(x_0, h) = f'(x_0, \frac{1}{\gamma+1} h_\gamma + \frac{\gamma}{\gamma+1}\tilde{h}) \leq \frac{1}{\gamma+1} f'(x_0, h_\gamma) + \frac{\gamma}{\gamma+1} f'(x_0, \tilde{h}) \leq \underbrace{\frac{\gamma}{\gamma+1}}_{>0} f'(x_0, \tilde{h}) < 0.$$

Portanto, $f'(x_0, h) < 0$, ou seja, $h \in D(f; x_0)$. Donde $F(Q; x_0) \subset D(f; x_0)$. □

Corolário 3.1. *Suponha que qualquer das seguintes condições se satisfaz:*

(i) *E é um espaço de Banach, $f(x)$ satisfaz uma condição de Lipschitz numa vizinhança do ponto x_0, $f(x)$ é diferenciável no ponto x_0 em qualquer direção, $f'(x, h)$ é convexa em h, e existe $\tilde{h}$ tal que $f'(x_0, \tilde{h}) < 0$;*

(ii) *$f(x)$ é um funcional convexo contínuo e existe $\tilde{x}$ tal que $f(\tilde{x}) < f(x_0)$;*

(iii) *E é um espaço de Banach, $f(x)$ é Fréchet-diferenciável em x_0 e $f'(x_0) \neq 0$.*

Então

$$F(Q; x_0) = D(f; x_0) = \{h : f'(x_0, h) < 0\}.$$

Teorema 3.6. *Se Q é um conjunto convexo, então o cone de direções factíveis no ponto $x_0 \in E$ vem dado por:*

$$F(Q; x_0) = \{\lambda(int(Q) - x_0), \lambda > 0\} = \{h : h = \lambda(x - x_0); \lambda > 0, x \in int(Q)\}.$$

Demonstração. Seja $A = \{h : h = \lambda(x - x_0); \lambda > 0, x \in int(Q)\}$. Se $h \in F(Q, x_0)$, então existe $\varepsilon_0 > 0$ e uma vizinhança U de h tal que $x_0 + \varepsilon\tilde{h} \in Q$ para todo $\varepsilon \in]0, \varepsilon_0[$ e para qualquer $\tilde{h} \in U$. Note que $V := \{x_0\} + \varepsilon U$ é uma vizinhança de $x_0 + \varepsilon h$ em Q, logo $x_0 + \varepsilon h \in int(Q)$. Assim $h = \dfrac{1}{\varepsilon}(x - x_0)$, onde $x = x_0 + \varepsilon h$ e $1/\varepsilon > 0$.
Por tanto, $h \in A$, ou seja, $F(Q, x_0) \subset A$.

Por outro lado, se $h \in A$, então $h = \lambda(x - x_0)$ com $\lambda > 0$ e $x \in int(Q)$. Como $x \in int(Q)$, então existe uma vizinhança $V \subset Q$ de x. Considere $U := \{\lambda(v - x_0); \lambda >$

$0, v \in V\}$ e $\varepsilon_0 = \frac{1}{\lambda}$ e note que U é uma vizinhança de h, logo para todo $\tilde{h} \in U$ e $0 < \varepsilon < \frac{1}{\lambda}$,

$$\begin{aligned} x_0 + \varepsilon\tilde{h} &= x_0 + \varepsilon(\lambda(v - x_0)); \; v \in V \subset Q \\ &= \lambda\varepsilon v + (1 - \lambda\varepsilon)x_0 \in Q, \end{aligned}$$

já que Q é convexo e $0 < \lambda\varepsilon = \frac{1}{\varepsilon_0}\varepsilon < 1$. Portanto, $h \in F(Q; x_0)$, ou seja, $A \subset F(Q; x_0)$. $\square$

3.2.3 Direções Tangentes

Quando um conjunto é determinado por certo operador diferenciável, o Teorema de Lyusternik (Teorema 3.10) é uma ferramenta muito forte para determinar o cone das direções tangentes. Antes de enunciar e demonstrar esse teorema, vejamos algumas observações e resultados.

Observação 3.6. *Das igualdades (3.1) e (3.2) da definição 3.9 da derivada de Fréchet, podemos formalizar tal definição em termos de δ e ε como segue.*
Dado $\varepsilon > 0$ arbitrário, existe $\delta > 0$ tal que

$$\|F(x_0 + h) - F(x_0) - \Lambda h + r(x_0, h)\| \leq \varepsilon \|h\| \tag{3.7}$$

para todo $h \in E_1$ e $\|h\| < \delta$.
Este último, naturalmente nos leva à seguinte definição.

Definição 3.10. *Seja $F : E_1 \longrightarrow E_2$ um operador, sendo E_1, E_2 espaços de Banach. Dizemos que F é estritamente diferenciável em $x_0 \in E_1$, se existe um operador linear $\Lambda : E_1 \longrightarrow E_2$ tal que para todo $\varepsilon > 0$, existe $\delta > 0$ e para todos $x_1, x_2 \in E_1$ satisfazendo $\|x_1 - x_0\| < \delta$ e $\|x_2 - x_0\| < \delta$ tem-se:*

$$\|F(x_1) - F(x_2) - \Lambda(x_1 - x_2)\| \leq \varepsilon \|x_1 - x_2\|. \tag{3.8}$$

Observação 3.7. *Tomando $x_2 = x_0$ e $x_1 = x_0 + h$ em (3.12), tem-se que F é diferenciável. Portanto toda função estritamente diferenciável é diferenciável, embora a recíproca não seja necessariamente é verdadeira.*

Os resultados a seguir, sem demostração, podem ser vistos em Alekseev [1], Ioffe [26] ou Flett [20].

Teorema 3.7. *Sejam E_1 e E_2 espaços de normados, $x_1, x_2 \in E$ e seja U um aberto de E_1 contendo $[x_1, x_2]$. Se $F : U \subset E_1 \longrightarrow E_2$ é diferenciável no ponto $x_0 \in [x_1, x_2]$, então*

$$\|F(x_1) - F(x_2) - \Lambda(x_1 - x_2)\| \leq \sup_{x_0 \in [x_1, x_2]} \|F'(x_0)\| \|x_1 - x_2\|. \tag{3.9}$$

Corolário 3.2. *Sob as condições no Teorema 6.5, tem-se*

$$\|F(x_1) - F(x_2) - \Lambda(x_1 - x_2)\| \leq \sup_{x_0 \in [x_1, x_2]} \|F'(x_0) - \Lambda\| \, \|x_1 - x_2\| \, .$$

Corolário 3.3. *Sejam E_1, E_2 espaços de Banach, U uma vizinhança de x_0 em E_1 e $F : U \subset E_1 \longrightarrow E_2$ é diferenciável, com $F'(x_0)$ contínua no ponto x_0. Então F é estritamente diferenciável em x_0.*

Teorema 3.8. *(**Existência das Funções Implícitas**) Sejam X um espaço topológico, Y e Z espaços de Banach, W uma vizinhança do ponto (x_0, y_0) em $X \times Y$. Considere uma aplicação $\Psi : W \to Z$ e $\Psi(x_0, y_0) = z_0$. Se as seguintes condições são verdadeiras:*

1) A aplicação $x \mapsto \Psi(x, y_0)$ é contínua em x_0;

2) Existe uma aplicação $\Lambda \in L(Y, Z)$ tal que $\forall \varepsilon > 0$, $\exists \delta > 0$ e uma vizinhança $\sum$ do ponto x_0, possuindo a propriedade a seguir:

- *A condição $x \in \sum$ e as desigualdades*

$$\|y_1 - y_0\| < \delta \;\; \|y_2 - y_0\| < \delta$$

implicam a desigualdade

$$\|\Psi(x, y_1) - \Psi(x, y_2) - \Lambda(y_1 - y_2)\| < \varepsilon \|y_1 - y_2\|;$$

3) $\Lambda Y \equiv Z$;

então existe um número $\kappa > 0$, uma vizinhança U do ponto (x_0, z_0) em $X \times Z$ e uma aplicação $\varphi : U \to Y$ tais que:

a) $\Psi(x, \varphi(x, z)) = z$;

b) $\|\varphi(x, z) - y_0\| \leq \kappa \|\Psi(x, y_0) - z\|$.

Teorema 3.9. *(Lyusternik). Seja $P : E_1 \to E_2$ um operador diferenciável numa vizinhança de x_0 e suponha que $P'(x_0) : E_1 \to E_2$ é contínua em uma vizinhança de x_0, além disso $P'(x_0)$ seja sobrejetivo. Então o conjunto das direções Tangente $T(Q; x_0)$ ao conjunto $Q = \{x : P(x) = 0\}$ no ponto x_0 é um subespaço de E, e esta dado por:*

$$T(Q; x_0) = \{h \in E : P'(x_0)h = 0\} := Ker P'(x_0).$$

Demonstração. Primeiramente suponhamos $h \in T^+(Q;x_0)$. Então, existem $\varepsilon > 0$ e $r : [0,\varepsilon] \to E_1$ tal que $x_0 + \alpha h + r(\alpha) \in Q$, com $\|r(\alpha)\| = o(\alpha)$, para $\alpha \to 0^+$ e como $P'(x_0)$ é contínua, logo, em virtude da diferenciabilidade estrita de P(Corolário 3.2) de acordo com a definição 3.10, temos que

$$0 = P(x_0 + \alpha h + r(\alpha)) = P(x_0) + \alpha P'(x_0)h + o(\alpha) = \alpha P'(x_0)h + o(\alpha).$$

Logo, $P'(x_0)h = 0$, donde $h \in KerP'(x_0)$ e consequentemente $T^+(Q;x_0) \subset KerP'(x_0)$.
Agora, para provarmos a outra inclusão, aplicaremos o Teorema das Funções Implícitas 3.8 à função $\Psi(x,h) = P(x+h)$, com $h_0 = 0$ e $z_0 = 0$.
Decorre da diferenciabilidade estrita de P que a aplicação $x \to \Psi(x,0)$ é contínua no ponto x_0, logo a condição 1) d e que

$$\begin{array}{rcl} \|P(x+\tilde{h}) - P(x+\hat{h}) - P'(x_0)(\tilde{h}-\hat{h})\| & \leq & \varepsilon\|\tilde{h}-\hat{h}\| \quad \Leftrightarrow \\ \|\Psi(x,\tilde{h}) - \Psi(x,\hat{h}) - P'(x_0)(\tilde{h}-\hat{h})\| & \leq & \varepsilon\|\tilde{h}-\hat{h}\|, \end{array}$$

quando $x \in B(x_0,\delta)$ e $\tilde{h}$, $\hat{h} \in B(0,\delta)$, donde 2) do é satisfeita. A condição 3) também é satisfeita, já que $P'(x_0)E_1 \equiv E_2$. Portanto teremos a existência de uma aplicação $\tilde{\varphi} : V \to E_1$, onde V é uma vizinhança do ponto x_0 em E_1 e $\tilde{\varphi}(x) = \varphi(x,z), \forall z \in U$, vizinhança de $(x_0,z_0) = (x_0,0)$. Considerando em particular para $z = 0$, segue que

$$\Psi(x,\varphi(x,z)) = \Psi(x,\tilde{\varphi}(x)) = 0 \Leftrightarrow P(x+\tilde{\varphi}(x)) = 0; \tag{3.10}$$

$$\begin{array}{rcl} \|\varphi(x,0) - h_0\| & = & \|\tilde{\varphi}(x)\| \leq \kappa\|\Psi(x,h_0) - z\| \\ & = & \kappa\|\Psi(x,0)\| = \kappa\|P(x)\| \end{array}$$

Ou seja:

$$\|\varphi(x,0) - h_0\| = \kappa\|P(x)\| \tag{3.11}$$

Definamos $r : [-\tilde{\varepsilon},\tilde{\varepsilon}] \to E_1$ tal que $r(\alpha) = \tilde{\varphi}(x_0 + \alpha h)$. Daí,

$$P(x_0 + \alpha h + r(\alpha)) = P(x_0 + \alpha h + \tilde{\varphi}(x_0 + \alpha h)) = 0,$$

por (3.10); de onde segue a afirmação $x_0 + \alpha h + r(\alpha) \in Q$. Temos também que:

$$\begin{array}{rcl} \|r(\alpha)\| & = & \|\tilde{\varphi}(x_0+\alpha h)\|, \text{ por (3.11)} \\ \|\tilde{\varphi}(x_0+\alpha h)\| & \leq & \kappa\|P(x_0+\alpha h)\| = \kappa\|P(x_0+\alpha h) - P(x_0)\| = \\ \kappa\|P'(x_0)\alpha h + o(\|\alpha h\|)\| & = & \kappa\|\alpha P'(x_0)h + o(\alpha)\|, \text{ para } \kappa > 0 \end{array}$$

Se $h \in kerP'(x_0)$, então $\|r(\alpha)\| = o(\alpha)$. Assim, pela definição de cone tangente, resulta que $h \in T(Q;x_0) \subseteq T^+_{x_0}Q$. Portanto, $kerP'(x_0) \subseteq T_{x_0}Q \subseteq T^+_{x_0}Q$. Desta maneira mostramos as igualdades:

$$T(Q;x_0) = T^+(Q;x_0) = kerP'(x_0) = \{h \in E_1; P'(x_0)h = 0\}.$$

$\square$

Observamos que o Teorema de Lyusternik é uma ferramenta poderosa para se obter o cone de direções tangentes, o qual é feito basicamente, pelo cálculo da derivada estrita do funcional associado. Porém, não é possível utilizá-lo no caso em que o operador associado não é estritamente diferenciável, ou quando a sua derivada não é sobrejetora. Na prática, a dificuldade se concentra na condição de sobrejetividade do operador derivada. Na terminologia de otimização, esta é uma condição de regularidade.

Exemplo 3.4. *Sejam* $E_1 = C([0,T];\mathbb{R}^n) \times L_\infty([0,T];\mathbb{R}^r)$, $x(t) \in C([0,T];\mathbb{R}^n)$, $u(t) \in L_\infty([0,T];\mathbb{R}^r)$, *e considere*

$$Q = \left\{ (x,u) \in E_1 : x(t) = c + \int_0^t \varphi\left(x(\tau), u(\tau), \tau\right) d\tau \right\}.$$

Defina um operador

$$P : E_1 \longrightarrow C([0,T];\mathbb{R}^n)$$

como

$$P(x,u) = x(t) - c - \int_0^t \varphi(x(\tau), u(\tau), \tau) d\tau,$$

com $\varphi_x(x,u,t)$ *e* $\varphi_u(x,u,t)$ *contínuas em* x, u, *mensuráveis em* t *e limitadas, para todo* $(x,u) \in E_1$ *limitadas e para todo* $t \in [0,T]$. *Então*

$$P'(x_0,u_0)(\bar{x},\bar{u}) = \bar{x}(t) - \int_0^t \left[\varphi_x(x_0(\tau), u_0(\tau), \tau)\bar{x}(\tau) + \varphi_u(x_0(\tau), u_0(\tau), \tau)\bar{u}(\tau)\right] d\tau \ \ e$$

$$T(Q,(x_0,u_0)) = \left\{(\bar{x},\bar{u}) : P'(x_0,u_0)(\bar{x},\bar{u}) = 0\right\}.$$

De fato, procedendo como no Exemplo 3.3 temos

$P(x+\bar{x}, u+\bar{u}) - P(x,u)$

$$\begin{aligned} &= \int_0^t \left[\varphi(x(\tau)+\bar{x}(\tau), u(\tau)+\bar{u}(\tau), \tau) - \varphi(x(\tau), u(\tau), \tau)\right] d\tau \\ &= \bar{x}(t) - \int_0^t \left[\varphi_x(x(\tau), u(\tau), \tau)\bar{x}(\tau) + \varphi_u(x(\tau), u(\tau), \tau), \bar{u}(\tau)\right] d\tau + \delta, \end{aligned} \tag{3.12}$$

onde $\delta := \int_0^t o\left(|(\bar{x}(\tau), \bar{u}(\tau))|\right) d\tau$ é um termo restante para os quais conseguimos a estimativa:

$$\delta = o\left(\sqrt{\|\bar{x}\|_C^2 + \|\bar{u}\|_{L_\infty}^2}\right).$$

De fato,

$$|(\bar{x}(\tau), \bar{u}(\tau))| = \sqrt{|\bar{x}(\tau)|^2 + |\bar{u}(\tau)|^2} \leq \sqrt{\|\bar{x}\|_C^2 + \|\bar{u}\|_{L_\infty}^2} = \|(\bar{x}, \bar{u})\|_{C \times L_\infty}, \tag{3.13}$$

com $C \times L_\infty := C([0,T]; \mathbb{R}^n) \times L_\infty([0,T]; \mathbb{R}^r)$,
isto implica que,

$$\frac{o(|(\bar{x}(\tau), \bar{u}(\tau))|)}{\|(\bar{x}, \bar{u})\|_{C \times L_\infty}} \leq \frac{o(|(\bar{x}(\tau), \bar{u}(\tau))|)}{|(x(\tau), u(\tau)|}. \tag{3.14}$$

Além disso, em (3.13) $\|(\bar{x}, \bar{u})\|_{C \times L_\infty} \to 0$ implica que $|(\bar{x}(\tau), \bar{u}(\tau))| \to 0$. Portanto, como $\varphi(x, u, t)$ é diferenciável em (x, u),temos que

$$\frac{o(|(\bar{x}(\tau), \bar{u}(\tau))|)}{\|(\bar{x}, \bar{u})\|} \to 0.$$

Logo, tendo em conta (3.14) temos

$$\frac{\int_0^t o(|(\bar{x}(\tau), \bar{u}(\tau))|)}{\|(\bar{x}, \bar{u})\|_{C \times L_\infty}} d\tau = \int_0^t \frac{o(|(\bar{x}(\tau), \bar{u}(\tau))|)}{\|(\bar{x}, \bar{u})\|_{C \times L_\infty}} d\tau \to 0.$$

Portanto, $\delta = o(\|(\bar{x}, \bar{u})\|_{C \times L_\infty})$, e note que

$$(\bar{x}, \bar{u}) \mapsto \bar{x}(t) - \int_0^t [\varphi_x(x(\tau), u(\tau), \tau)\bar{x}(\tau) + \varphi_u(x(\tau), u(\tau), \tau)\bar{u}(\tau)] \, d\tau$$

é um operador linear de E_1 em $C([0,T]; \mathbb{R}^n)$. Então segue que $P(x, u)$ é diferenciável e

$$P'(x, u)(\bar{x}, \bar{u}) = \bar{x}(t) - \int_0^t [\varphi_x(x(\tau), u(\tau), \tau)\bar{x}(\tau) + \varphi_u(x(\tau), u(\tau), \tau)\bar{u}(\tau)] \, d\tau.$$

$P'(x, u)$ é contínua em alguma vizinhança de (x, u), desde que $\varphi_x(x, u, t)$ e $\varphi_u(x, u, t)$ são também contínuas.
Vejamos agora que $P'(x, u)$ é sobrejetora. Em outras palavras, vamos mostrar que para todo $a(t) \in C([0,T]; \mathbb{R}^n)$, existe $(\bar{x}(t), \bar{u}(t)) \in E_1$ tal que

$$a(t) = \bar{x}(t) - \int_0^t [\varphi_x(x_0(\tau), u_0(\tau), \tau)\bar{x}(\tau) + \varphi_u(x_0(\tau), u_0(\tau), \tau)\bar{u}(\tau)] \, d\tau,$$

ou seja, $P(x, u)(\bar{x}, \bar{u}) = a(t)$. Tomando $\bar{u}(t) = 0$, temos que

$$\bar{x}(t) = a(t) + \int_0^t \varphi_x(x_0(\tau), u_0(\tau), \tau)\bar{x}(\tau) d\tau \tag{3.15}$$

é uma equação integral de Volterra de segunda espécie, e a teoria de equações integrais nos diz que para todo $a(t) \in C([0,T]; \mathbb{R}^n)$, (3.15) possui uma única solução $\bar{x}(t) \in C([0,T]; \mathbb{R}^n)$. Para mais detalhes veja Functional Analysis, Kantorovich L.V. and Akilov G.P. [[27] p 396].

Portanto, dado $a(t) \in C([0,T];\mathbb{R}^n)$, existe $(\bar{x},\bar{u}) \in E_1$ tal que $P(x,u)((\bar{x},\bar{u})) = a(t)$, consequentemente, $P'(x,u)$ é sobrejetora. Logo do Teorema de Lyusternik (Teorema 3.9) segue que

$$T(Q;(x,u)) = \{(\bar{x},\bar{u}) : P'(x,u)(\bar{x},\bar{u}) = 0\}.$$

3.2.4 Cálculo de Cones Duais

Teorema 3.10. *Seja K um subespaço de E. Então $K^* = \{f \in E' : f(x) = 0 \ \forall x \in K\}$.*

Demonstração. Dado $f(x) > 0$, para algum $x \in K$ e $f \in K^*$. Como $-x \in K$, desde que K é subespaço vetorial, então $f(-x) = -f(x) < 0$ o qual contradiz o fato de $f \in K^*$. □

Teorema 3.11. *Seja $f \in E'$ fixado, Então:*

i) Se $K_1 = \{x \in E : f(x) = 0\}$, então $K_1^* = \{\lambda f : \lambda \in -\infty < \lambda < \infty\}$;

ii) Se $K_2 = \{x \in E : f(x) \geq 0\}$, então $K_2^* = \{\lambda f : \lambda \geq 0\}$;

iii) Se $K_3 = \{x \in E : f(x) > 0\}$, então $K_3^* = E'$ se $f = 0$, e $K_3^* = K_2^*$ se $f \neq 0$.

Demonstração. i) Se $g \in K^*$, pelo teorema anterior, $g(x) = 0 \ \forall x \in K_1$, pelo lema (2.3) segue que $g = \lambda f$ com $\lambda \in \mathbb{R}$. Assim, $K_1^* = \{\lambda f : \lambda \in -\infty < \lambda < \infty\}$.
Por outro lado, se $g \in \{\lambda f : \lambda \in -\infty < \lambda < \infty\}$, então $g = \lambda f \in K$, mas $f(x) = 0, \forall x \in K$, daí que $g(x) = 0$, $\forall x \in K$. Portanto $g \in K^*$. Assim mostramos a igualdade.

ii) Como $K_1 \subset K_2$, segue do Lema 5.5 $K_2^* \subset K_2^*$. Agora se $g = \lambda f \in K_2^*$, temos que $g(x) = \lambda f(x) \geq 0$, para todo $x \in K_2$. Mas $f(x) \geq 0$, então $\lambda \geq 0$.

iii) Se f=0, então $K_3 = \emptyset$ e portanto $K^* = E'$. Mas, se $f \neq 0$, temos que $K_2 = K_3$, isto implica que $K_2^* = K_3^*$. □

Definição 3.11. *O conjunto de* ***funcionais suporte*** *para um conjunto Q no ponto x_0 é definido por $Q^* = \{f \in E' : f(x) \geq f(x_0), \forall x \in Q\}$.*

Em alguns casos os duais dos cones de direções factíveis e de direções tangentes coincidem com o conjunto de funcionais suporte para Q no ponto x_0.

Teorema 3.12. *Seja Q um conjunto convexo e fechado. Então $T^*(Q;x_0) = Q^*$. Se além disso $int(Q) \neq \emptyset$, então $F^*(Q;x_0) = Q^*$.*

Demonstração. Sejam dados um funcional $f \in Q^*$ e uma direção tangente unilateral $h \in T^+(Q;x_0)$. Então dado $0 < \varepsilon < \varepsilon_0$, existe $x_\varepsilon \in Q$, $x_\varepsilon = x_0 + \varepsilon h + r(\varepsilon)$, tal que $\frac{r(\varepsilon)}{\varepsilon} \in V$, para toda vizinhança V de zero e ε suficientemente pequeno. Já que $x_\varepsilon \in Q$ e $f \in Q^*$, temos que $f(x_\varepsilon) \geq f(x_0)$. Portanto

$$f(h) = \frac{f(x_\varepsilon) - f(x_0)}{\varepsilon} - \frac{f(r(\varepsilon))}{\varepsilon} \geq -f\left(\frac{r(\varepsilon)}{\varepsilon}\right).$$

Mas $f\left(\frac{r(\varepsilon)}{\varepsilon}\right) \to 0$, quando $\varepsilon \to 0^+$ (pois, desde que f é contínua, para cada $\delta > 0$, existe uma vizinhança V de 0, tal que para qualquer $x \in V$, $|f(x)| \leq \delta$. Assim, $\left|f\left(\frac{r(\varepsilon)}{\varepsilon}\right)\right| < \delta$, para $\varepsilon < 0$ suficientemente pequeno). Por conseguinte, $f(h) \geq 0$, isto é, $f \in T(Q, x_0)^*$.

Reciprocamente, sejam $f \in T(Q, x_0)^*$, $x \in Q$. Então $h = x - x_0$ é uma direção tangente (para $0 < \varepsilon < 1$, tome $x_0 + \varepsilon h \in E_1$, donde obtêm-se que $x_0 + \varepsilon h = (1-\varepsilon)x_0 + \varepsilon x \in C$, já que Q é convexo). Visto que $f \in T(Q, x_0)^*$, segue que $f(h) \geq 0$, ou seja, $f(x) \geq f(x_0)$. Isto significa que $f \in Q^*$.
Deste modo, provamos que $Q^* = T(Q;x_0)^*$.

Se $intQ \neq \emptyset$, então pelo Teorema 3.6,

$$F(Q;x_0) = \{h \in E_1 : h = \lambda(x - x_0), x \in intQ \text{ e } \lambda > 0\}.$$

Consequentemente, se $f \in F(Q;x_0)^*$, então, para todo $x \in intQ$ e todo $\lambda \geq 0$,

$$f(\lambda(x - x_0)) \geq 0 \text{ equivalentemente, } \lambda f(x) \geq \lambda f(x_0).$$

Em particular, para $\lambda = 1$. Entretanto, $cl(intQ) = clQ$, quando $intQ \neq \emptyset$. Portanto, pela continuidade do funcional f, o fato de que $f(x) \geq f(x_0)$, para todo $x \in intQ$, implica que $f(x) \geq f(x_0)$, para todo $x \in clQ$. Podemos então dizer que $f \in Q^*$ e $F(Q;x_0)^* \subseteq T(Q;x_0)^* = Q^*$. Por outro lado, $F(Q;x_0) \subseteq T(Q;x_0)$, por conseguinte do Lema 2.1, $F(Q;x_0)^* \supset T(Q;x_0)^* = Q^*$. Logo, $F(Q;x_0)^* = Q^*$. □

Observação 3.8. *Do resultado acima, verifica-se que a determinação dos cones duais é equivalente a determinação dos funcionais suporte para um conjunto convexo e fechado em E_1.*

Lema 3.6. *Seja $U = \{u \in L_\infty([0,T];\mathbb{R}^r) : u(t) \in M$ q.t.p. $0 \leq t \leq T\}$, $M \subset \mathbb{R}^r$ e $u^\circ \in U$. Então, se o funcional definido por*

$$f(u) = \int_0^T \langle g(t), u(t)\rangle dt$$

é o suporte para U no ponto u_0, com $g(t) \in L_1([0,T];\mathbb{R}^r)$, então $\langle g(t), (u - u^\circ)(t)\rangle \geq 0$

para todo $u(t) \in M$ e q.t.p. $t \in [0,T]$(isto é, para quase todo $t \in [0,T]$ o vetor $g(t) \in \mathbb{R}^r$ é um suporte a M no ponto $u^o(t)$).

Demonstração. Suponha que a afirmação é falsa, ou seja, existe um subconjunto $R_1 \subset [0,T]$ com $\mu(R_1) \neq 0$ (onde μ denota a medida de Lebesgue em $\mathbb{R}$), tal que, para todo $t \in R_1$, existe $\bar{u}(t) \in M$ e $\langle g(t), (\bar{u} - u^\circ)(t)\rangle < 0$. Pelo Teorema de Lusin (veja Apêndice D), dado um intervalo J e $\phi : J \longrightarrow \mathbb{R}^r$ uma função mensurável, então, dado $\epsilon > 0$, existe um subconjunto $J_\epsilon \subset J$ tal que $\mu(J \setminus J_\epsilon) \leq \epsilon$ e ϕ é contínua em J_ϵ. Logo existem subconjuntos $R_2 \subset [0,T]$, com $\mu(R_2) < \frac{\epsilon}{2}$, $R_3 \subset [0,T]$,com $\mu(R_3) < \frac{\epsilon}{2}$ tais que g é contínua sobre $[0,T] \setminus R_2$ e u° é contínua $[0,T] \setminus R_3$.

Como $\mu(R_2) + \mu(R_3) < \mu(R_1)$, existe um ponto $t_0 \in R_1$, $t_0 \notin R_2 \cup R_3$. Agora, como $g(t)$ e $u^o(t)$ são contínuas em t_0 e $\langle g(t_0), \bar{u}(t_0) - u^o(t_0)\rangle = \gamma < 0$, existe $R_4 \subset [0,T]$ com $\mu(R_4) > 0$, tal que $\langle g(t), \bar{u}(t_0) - u^o(t)\rangle \leq \frac{\gamma}{2}$ para $t \in R_4$. Considere a função

$$u_1(t) = \begin{cases} \bar{u}(t_0), & \text{se } t \in R_4, \\ u^o(t), & \text{se } t \in [0,1] \setminus R_4. \end{cases}$$

Então, $u_1 \in U$ ao mesmo tempo que

$$f(u_1) = \int_0^T \langle g(t), u^o(t)\rangle dt + \int_{R_4} \langle g(t), \bar{u}(t_0) - u^o(t)\rangle dt \leq f(u^o) + \frac{\gamma}{2}\mu(R_4),$$

ou seja, $f(u_1) < f(u^o)$, o que contradiz o fato de f ser o funcional suporte a U em u^o. $\square$

CAPÍTULO 4

O PRINCÍPIO DO MÁXIMO DE PONTRYAGIN LOCAL

Neste capítulo estudamos um problema básico de Controle Ótimo em tempo contínuo. O resultado fundamental que dá condições necessárias de otimalidade é o Princípio do Máximo de Pontryagin cuja demostração será feita usando o formalismo de Dubovistkii e Milyutin estudado no Capítulo [3].

4.1 Problema de Controle Ótimo Princípio do Máximo Local

Apresentamos uma versão simplificada de um problema de controle ótimo, logo verifica-se que se enquadra na formulação de problema extremal e aplicamos a condição de extremo, Teorema 3.1, para obter o Princípio do Máximo de Pontryagin Local.

Problema 4.1. *Encontrar funções $x : [0, T] \to \mathbb{R}^n$, (trajetória de fase) e $u : [0, T] \to \mathbb{R}^r$ (controle) que satisfaçam a equação diferencial*

$$\frac{dx(t)}{dt} = \varphi(x(t), u(t), t) \;\; q.t.p. \;\; 0 \leq t \leq T, \tag{4.1}$$

com as condições de fronteira

$$x(0) = c, \tag{4.2}$$

$$x(T) = d, \tag{4.3}$$

sendo $\varphi : \mathbb{R}^n \times \mathbb{R}^r \times [0,T] \to \mathbb{R}^n$, *de modo que minimizem o funcional integral*

$$\int_0^T \Phi(x(t), u(t), t)dt, \tag{4.4}$$

cujo controle satisfaz restrições do tipo

$$u(t) \in M, \; q.t.p. \; 0 \leq t \leq T, \tag{4.5}$$

com $\Phi : \mathbb{R}^n \times \mathbb{R}^r \times [0,T] \to \mathbb{R}$, $M \subset \mathbb{R}^r$.

Como conjunto de controles admissíveis tomamos a classe de todas as funções mensuráveis limitadas (isto é, $u \in L_\infty([0,T]; \mathbb{R}^r)$*). Além disso, em vez da equação (4.1) consideramos a equação integral equivalente, ou seja, a solução será o par*

$$(x, u) \in C([0,T]; \mathbb{R}^n) \times L_\infty([0,T]; \mathbb{R}^r)$$

satisfazendo a equação integral

$$x(t) = c + \int_0^t \varphi(x(\tau), u(\tau), \tau)d\tau \; q.t.p. \; 0 \leq t \leq T. \tag{4.6}$$

Teorema 4.1. *(Princípio do Máximo Local de Pontryagin. Sejam* $\varphi : \mathbb{R}^n \times \mathbb{R}^r \times [0,T] \to \mathbb{R}^n$, $\Phi : \mathbb{R}^n \times \mathbb{R}^r \times [0,T] \to \mathbb{R}$, *tais que* $\varphi(x(t), u(t), t)$ *e* $\Phi(x(t), u(t), t)$ *são funções contínuas em* x *e* u; *mensuráveis em* t; *continuamente diferenciáveis com respeito a* x *e* u. *Além disso, sejam* $\varphi_x(x(t), u(t), t)$, $\varphi_u(x(t), u(t), t)$, $\Phi_x(x(t), u(t), t)$, $\Phi_u(x(t), u(t), t)$ *funções limitadas para todas* x *e* u *limitadas e* $M \subset \mathbb{R}^r$ *convexo e fechado, tal que* $intM \neq \emptyset$. *Se* $(x^o(t), u^o(t))$ *é uma solução do Problema 4.1. Então existem um número* $\lambda_o \geq 0$ *e uma função* $\psi : [0,T] \to \mathbb{R}^n$, *satisfazendo a equação*

$$\frac{d\psi(t)}{dt} = -\varphi_x^*(x^o(t), u^o(t), t)\psi(t) + \lambda_o \Phi_x(x^o(t), u^o(t), t) q.t.p. \; 0 \leq t \leq T, \tag{4.7}$$

tais que λ_o *e* $\psi(.)$ *não são simultaneamente iguais (identicamente) a zero, além disso*

$$\langle -\varphi_u^*(x^o(t), u^o(t), t)\psi(t) + \lambda_o \Phi_u(x^o(t), u^o(t), t), u(t) - u^o(t) \rangle \geq 0 \tag{4.8}$$

para quase todo $0 \leq t \leq T$ *e para todo* $u \in M$.

Observação 4.1. *Aqui, o espaço* E *é tomado como*

$$E = C([0,T]; \mathbb{R}^r) \times L_\infty([0,T]; \mathbb{R}^r).$$

Denotamos por

$$\begin{aligned} Q_1 &= \{(x,u) \in E : \text{ é satisfeita } (4.5)\} \text{ e} \\ Q_2 &= \{(x,u) \in E : \text{são satisfeitas } (4.1), (4.2) \text{ e } (4.3)\}. \end{aligned}$$

Então o Problema 4.1 torna-se

$$\begin{cases} \min F(x,u) = \displaystyle\int_0^T \Phi(x(t), u(t), t)dt \\ s.a. \\ x \in Q = Q_1 \bigcap Q_2. \end{cases}$$

Logo procede-se o esquema geral, onde Q_1 é conjunto das restrições de desigualdade, Q_2 o conjunto das restrições de igualdade, após se determinam os cones e seu respetivos duais. Finalmente aplica-se o teorema de Dubovitskii-Milyutin.

Demonstração. Analisaremos o problema de acordo com o esquema geral. Primeiramente iremos determinar os cones associados ao funcional F, aos conjuntos Q_1 e Q_2, e por fim calcularemos os cones duais.

a) Análise do funcional F.

Como $F(x,u)$ é diferenciável (ver Exemplo 3.3) e dado $(\bar{x}, \bar{u}) \in E$,

$$F'(x^o, u^o)(\bar{x}, \bar{u}) = \int_0^T \left[\langle \Phi_x(x^o(t), u^o(t), t), \overline{x}(t)\rangle + \langle \Phi_u(x^o(t), u^o(t), t), \bar{u}(t)\rangle\right] dt,$$

do Teorema 3.3, o cone de direções de descida é dado por

$$D(F; (x^o, u^o)) = \{(\bar{x}, \bar{u}) \in E : \ F'(x^o, u^o)(\bar{x}, \bar{u}) < 0\},$$

e do Teorema 3.10, se $D(F; (x^o, u^o)) \neq \emptyset$, então para qualquer $f_0 \in D(F; (x^o, u^o))^*$,

$$f_0(\overline{x}, \overline{u}) = -\lambda_o \int_0^T \left[\langle \Phi_x(x^o(t), u^o(t), t), \overline{x}(t)\rangle + \langle \Phi_u(x^o(t), u^o(t), t), \overline{u}(t)\rangle\right] dt, \qquad (4.9)$$

para algum $\lambda_o \geq 0$, ou seja,

$$D(F; (x^o, u^o))^* = \{-\lambda_0 F'(x^o, u^o) : \lambda_0 \geq 0\}.$$

b) Análise da restrição de desigualdade em Q_1.

O conjunto $Q_1' = \{u \in L_\infty([0,T]; \mathbb{R}^r) : u(t) \in M \subset \mathbb{R}^r \ q.t.p. \ 0 \leq t \leq T]\}$ é convexo e fechado no espaço $L_\infty([0,T]; \mathbb{R}^r)$ desde que M é convexo, fechado e $int(M) \neq \emptyset$. Portanto o conjunto $Q_1 = C([0,T];\mathbb{R}^n) \times Q_1'$ é também convexo e fechado em E, além disso

$$int(Q_1) = C([0,T];\mathbb{R}^n) \times int(Q_1') \neq \emptyset.$$

Seja $F(Q_1;(x^o,u^o))$ o cone de direções factíveis para Q_1 no ponto (x^o,u^o). Então, se $f_1 \in F(Q_1;(x^o,u^o))^*$, segue-se que $f_1 = (0, f_1')$, onde $f_1' \in L_\infty^*([0,T];\mathbb{R}^r)$ é um funcional suporte para Q_1' no ponto u^o, conforme Teorema 3.10 e Teorema 3.12.

c) Análise da restrição em Q_2.

Queremos calcular o subespaço tangente $T(Q_2;(x^o,u^o))$ ao conjunto

$$Q_2 = \left\{(x,u) \in E : x(t) = c + \int_0^t \varphi(x(\tau),u(\tau),\tau)d\tau,\ x(T) = d,\ 0 \leq t \leq T\right\}$$

no ponto (x^o,u^o). Para isto definimos o operador $\Lambda : E \longrightarrow E_1$, onde $E_1 = C([0,T];\mathbb{R}^n) \times \mathbb{R}^n$ por

$$\Lambda(x,u) = \left(x(t) - c - \int_0^t \varphi(x(\tau),u(\tau),\tau)d\tau, x(T) - d\right).$$

Logo

$$Q_2 = \{(x,u) \in E : \Lambda(x,u) = (0,0)\}.$$

Além disso,

$\Lambda(x^o + \bar{x}, u^o + \bar{u}) - \Lambda(x^o,u^o)$

$$\begin{aligned} &= \left(\bar{x}(t) - \int_0^t [\varphi((x^o+\bar{x})(\tau),(u^o+\bar{u})(\tau),\tau) - \varphi(x^o(\tau),u^\circ(\tau),\tau)]\,d\tau, \bar{x}(T)\right) \\ &= \left(\bar{x}(t) - \int_0^t [\varphi_x(x^o(\tau),u^o(\tau),\tau)\bar{x}(\tau) + \varphi_u(x^o(\tau),u^o(\tau),\tau)\bar{u}(\tau)]\,d\tau, \bar{x}(T)\right) + \delta. \end{aligned}$$

Onde para o resto δ pode-se encontrar uma estimativa:

$$\delta = o\left(\sqrt{\|\bar{x}\|_C^2 + \|\bar{u}\|_{L_\infty}^2}\right).$$

Note que

$$\left(\bar{x}(t) - \int_0^t [\varphi_x(x^o(\tau),u^o(\tau),\tau)\bar{x}(\tau) + \varphi_u(x^o(\tau),u^o(\tau),\tau)\bar{u}(\tau)]\,d\tau, \bar{x}(T)\right)$$

é um operador linear de E em E_1. Logo Λ é diferenciável em (x^o,u^o), como no Exemplo 3.3, e

$$\Lambda'(x^o,u^o)(\bar{x},\bar{u}) = \left(\bar{x}(t) - \int_0^t [\varphi_x(x^o(\tau),u^o(\tau),\tau)\bar{x}(\tau) + \varphi_u(x^o(\tau),u^o(\tau),\tau)\bar{u}(\tau)]\,d\tau, \bar{x}(T)\right).$$

Além disso, $\Lambda'(x^o,u^o)$ é contínua em qualquer vizinhança de $(\bar{x},\bar{u})$. Se mostrarmos que

$\Lambda'(x^o, u^o)$ é sobrejetor, então pelo Teorema de Lyusternik (Teorema 3.9), podemos exibir o cone tangente $T(Q_2, (x^o, u^o))$.

Afirmacção. *Se a condição de não-degeneração* ($\boldsymbol{A}$) *é satisfeita para o sistema*

$$\begin{cases} \dfrac{d\bar{x}(t)}{dt} = A(t)\bar{x}(t) + B(t)\bar{u}(t), \\ \bar{x}(0) = 0, \end{cases}$$

com $A(t) = \varphi_x(x^o(t), u^o(t), t)$ e $B(t) = \varphi_u(x^o(t), u^o(t), t)$, então $\Lambda'(x^o, u^o)$ é sobrejetora, ou seja, $\Lambda(x^o, u^o)'(E) = E_1$. Assim todas as hipóteses do Teorema de Lyusternik (Teorema 3.9) são satisfeitas para operador Λ, portanto o cone tangente (subespaço tangente) $T(Q_2, (x^o, u^o))$ é dado por

$$T(Q_2, (x^o, u^o)) = \{(\bar{x}, \bar{u}) \in E : \Lambda'(x^o, u^o)(\bar{x}, \bar{u}) = 0\},$$

isto é,

$$\frac{d\bar{x}(t)}{dt} = \varphi_x(x^o(t), u^o(t), t)\bar{x}(t) + \varphi_u(x^o(t), u^o(t), t)\bar{u}(t), \bar{x}(0) = 0, \tag{4.10}$$

$$\bar{x}(T) = 0. \tag{4.11}$$

Sejam $L_1 \subset E$ o conjunto dos pares $(\bar{x}, \bar{u})$ satisfazendo (4.10) e $L_2 \subset E$ o conjunto dos pares $(\bar{x}, \bar{u})$ satisfazendo (4.11). Então L_1 e L_2 são subespaços e $T(Q_2, (x^o, u^o)) = L_1 \cap L_2$, ou seja,

$$L_1 = \{(\bar{x}, \bar{u}) \in E : \bar{x}(\cdot) \text{ verifica } (4.10)\} \text{ e } L_2 = \{(\bar{x}, \bar{u}) \in E : \bar{x}(T) = 0\}.$$

Fato 1. Sendo L_1 subespaço, segue do Teorema 3.10 que se $f_3 \in L_1^*$, temos que $f_3(\bar{x}, \bar{u}) = 0$ para todo $(\bar{x}, \bar{u}) \in E$ satisfazendo (4.10).

Fato 2. Se $f_2 \in L_2^*$, então $f_2(\bar{x}, \bar{u}) = \langle \bar{x}(T), a\rangle$, onde $a \in \mathbb{R}^n$. De fato, como L_2 é subespaço vetorial, pelo Teorema 3.10, $f_2(\bar{x}, \bar{u}) = 0$ $\forall (\bar{x}, \bar{u}) \in L_2$. Portanto, $f_2(\bar{x}(T), \bar{u}(T)) = 0$ para todo $(\bar{x}, \bar{u}) \in L_2$, donde existe $(a, b) \in \mathbb{R}^n \times \mathbb{R}^r$ tal que

$$0 = f_2(\bar{x}(T), \bar{u}(T)) = \langle (a, b), (\bar{x}(T), \bar{u}(T))\rangle = \langle a, \bar{x}(T)\rangle + \langle b, \bar{u}(T)\rangle.$$

Claro que $b = 0$, portanto, se $f_2 \in L_2^*$, temos que $f_2(\bar{x}, \bar{u}) = (\bar{x}(T), a)$.
Considere

$$L_2 = \{(\bar{x}, \bar{u}) : \bar{x}(T) = 0\} = \bigcap_{i=1}^{n} L_{2i}$$

onde $L_{2i} = \{(\bar{x}, \bar{u}) \in E : \bar{x}_i(T)\}$, $i = 1, 2, \cdots, n$.
Daí segue que,

$$\sum_{i=1}^{n} L_{2i}^* = \{\lambda_1 \bar{x}_1(T) + \lambda_2 \bar{x}_2(T) + \cdots + \lambda_n \bar{x}_n(T); \lambda_i \in \mathbb{R}\} = \{(\bar{x}(T), \lambda) : \lambda \in \mathbb{R}^n\}.$$

Sendo L_{2i} fechado, $\forall i = 1, 2.., n$ (w-fechado), convexo e $\sum L_{2_1}^*$ w^*-fechado, temos que

$$L_2^* = \left(\bigcap_{i=1}^{n} L_{2i}\right)^* = \sum_{i=1}^{n} L_{2i}^* = \{(\bar{x}(T), a) : a \in \mathbb{R}^n\}.$$

Assim, L_2^* tem dimensão finita n e como a soma de subespaços fechados de dimensão finita é também fechado, $L_1^* + L_2^*$ é fechado, logo w^*-fechado (L_1^* e L_2^* são fechados). Pelo Teorema 2.4, segue que

$$T(Q_2, (x^o, u^o))^* = (L_1 \cap L_2)^* = L_1^* + L_2^*.$$

Portanto, se $f \in T(Q_2, (x^o, u^o))^* \Rightarrow f(\bar{x}, \bar{u}) = f_2(\bar{x}, \bar{u}) + f_3(\bar{x}, \bar{u}) = \langle \bar{x}(T), a \rangle$, $a \in \mathbb{R}^n$.
$\Rightarrow f_2(\bar{x}, \bar{u}) = \langle \bar{x}(T), a \rangle$.

d) Análise da Equação de Euler.

Aplicando o Teorema de Dubovitskii-Milyutin (Teorema 3.1), temos que existem $f_0, f_1, f_2, f_3 \in E'$ não todos nulos, tal que para todo $(\bar{x}, \bar{u}) \in E$,

$$f_0(\bar{x}, \bar{u}) + f_1(\bar{x}, \bar{u}) + f_2(\bar{x}, \bar{u}) + f_3(\bar{x}, \bar{u}) = 0, \tag{4.12}$$

onde, $f_0(\bar{x}, \bar{u})$ é dado por (4.9); $f_1(\bar{x}, \bar{u}) = f_1'0(\bar{u})$ é o suporte para Q_1' em u^0; $f_3(\bar{x}, \bar{u})$ se anula em $(\bar{x}, \bar{u})$, solução de (4.10) e $f_2(\bar{x}, \bar{u}) = \langle \bar{x}(T), a \rangle$, $a \in \mathbb{R}^n$.

e) Análise da Equação de Euler.

A equação (4.10) vale para todo $(\bar{x}, \bar{u}) \in E$. Tome $\bar{u}$ arbitrário e $\bar{x} = \bar{x}(\bar{u})$ uma solução de (4.10). Com essa escolha de $(\bar{x}, \bar{u})$ temos que $f_3(\bar{x}, \bar{u}) = 0$, e a equação (4.12) torna-se

$$f_1'(\bar{u}) = \lambda_0 \int_0^T [\langle \Phi_x(x^o, u^o, t), \bar{x} \rangle + \langle \Phi_u(x^o, u^o, t), \bar{u} \rangle] \, dt - \langle a, \bar{x}(T) \rangle. \tag{4.13}$$

Vamos trabalhar com o lado direito, de forma a substituir $\bar{x}$ por $\bar{u}$. Para isto consideremos $\psi(t)$ solução do sistema (4.7) com a condição inicial de fronteira $\psi(T) = a$.
Usando integração por partes e o fato que $(\bar{x}, \bar{u})$ satisfaz (4.7), temos

$$
\begin{aligned}
&\lambda_0 \int_0^T \langle \Phi_x(x^o(t), u^o(t), t), \overline{x}(t)\rangle \, dt - (a, \overline{x}(T)) \\
&= \int_0^T \left\langle \frac{d}{dt}\psi(t) + \Phi_x^*(x^o(t), u^o(t), t)\psi(t), \bar{x}(t) \right\rangle dt - \langle \bar{x}(T), a\rangle \\
&= \int_0^T \left\langle \frac{d}{dt}\psi(t), \bar{x}(t) \right\rangle dt + \int_0^T \langle \phi_x^*(x^o(t), u^o(t), t)\psi(t), \bar{x}(t)\rangle \, dt - \langle \bar{x}(T), a\rangle \\
&= \langle \psi(t), \bar{x}(t)\rangle \,|_0^T - \int_0^T \left\langle \psi(t), \frac{d}{dt}\bar{x}(t) \right\rangle dt + \int_0^T \langle \psi(t), \Phi_x(x^o, u^o, t)\bar{x}(t)\rangle \, dt - \langle \bar{x}(T), a\rangle \\
&= -\int_0^T \langle \psi(t), \Phi_u(x^o, u^o, t)\bar{u}\rangle \, dt \\
&= -\int_0^T \langle \Phi_u^*(x^o, u^o, t)\psi(t), \bar{u}\rangle \, dt.
\end{aligned}
$$

Daí que, (4.12) pode ser escrita como

$$
f_1'(\bar{u}) = \int_0^T \langle -\phi_u^*(x^o, u^o, t)\psi(t) + \lambda_0 \phi_u(x^o, u^o, t), \bar{u}\rangle \, dt,
$$

onde $\bar{u}$ é arbitrário e $f'(\bar{u})$ é o suporte para Q_1' no ponto u^o. Logo, por definição de suporte de Q_1' em u^o, temos que $f_1'(u) \geq f_1'(u^\circ)$ q.t.p. $0 \leq t \leq T$ e $\forall u \in M$. Portanto, do Lema 3.6 segue que
$\langle (-\phi_u^*(x^o, u^o, t)\psi(t) + \lambda_0 \phi_u(x^o, u^o, t), u - u^o(t)\rangle \geq 0$ q.t.p. $0 \leq t \leq T$ e $\forall u \in M$.
Logo, temos que (4.8) é satisfeita. Observe que sob estas hipóteses, o caso $\lambda_0 = 0$ e $\psi(t) = 0$ não acontecem simultaneamente pois, se ocorressem, teríamos $f_0(\bar{x}, \bar{u}) = 0$; $a = \psi(T) = 0$, então $f_1'(u) = 0$ e $f_2(\bar{x}, \bar{u}) = 0$. Logo teríamos $f_0 = f_1 = f_2 = f_3 = 0$ o que contradiz a condição de não-trivialidade do Teorema de Duvobitskii Milyutin (Teorema 3.3).

f) Analise de casos excepcionais.

No decorrer da demostração, fizemos duas hipóteses adicionais:

1) $D(F; (x^o, u^o)) \neq \emptyset$;

2) O sistema (4.10) é não-degenerado.

Mostremos que estas duas hipóteses são supérfluas. De fato:
1) Se $K_0 = \emptyset$, então

$$
\int_0^T [\langle \Phi_x(x^o, u^o, t), \bar{x}(t)\rangle + \langle \Phi_u(x^o, u^o, t), \bar{u}(t)\rangle] \, dt = 0 \ \forall (\bar{x}, \bar{u}) \in E. \tag{4.14}
$$

Tomando $\lambda_0 = 1; \psi(T) = 0$, temos:

$$\begin{aligned}
\int_0^T \langle \Phi_x(x^o,u^o,t), \bar{x}(t)\rangle\, dt &= \int_0^T \left\langle \frac{d}{dt}\psi(t) + \varphi_x^*(x^o,u^o,t)\psi(t), \bar{x}(t)\right\rangle dt \\
&= \int_0^T \left\langle \frac{d}{dt}\psi(t), \bar{x}(t)\right\rangle dt + \int_0^T \langle \varphi_x^*(x^o,u^o,t)\psi(t), \bar{x}(t)\rangle\, dt \\
&= \langle \psi(t), \bar{x}(t)\rangle\, |_0^T - \int_0^T \left\langle \psi(t), \frac{d\bar{x}(t)}{dt}\right\rangle dt \\
&\quad + \int_0^T \langle \psi(t), \varphi_x(x^o,u^o,t)\bar{x}\rangle\, dt.
\end{aligned}$$

Como $(\bar{x},\bar{u})$ satisfaz (4.10) temos que

$$\begin{aligned}
&= -\int_0^T \langle \psi(t), \varphi_x(x^o,u^o,t)\bar{x}(t)\rangle\, dt - \int_0^T \langle \psi(t), \varphi_u(x^o,u^o,t)\bar{u}(t)\rangle\, dt \\
&\quad + \int_0^T \langle \psi(t), \varphi_x(x^o,u^o,t)\bar{x}(t)\rangle\, dt \\
&= +\int_0^T \langle \varphi_u^*(x^o,u^o,t)\psi(t), \bar{u}(t)\rangle\, dt.
\end{aligned}$$

Portanto

$$\int_0^T \langle \Phi_x(x^o,u^o,t), \bar{x}(t)\rangle\, dt = -\int_0^T \langle \varphi_u^*(x^\circ,u^\circ,t)\psi(t), \bar{u}(t)\rangle\, dt. \tag{4.15}$$

Por (4.14) e (4.15) temos que:

$$\int_0^T \langle -\varphi_u^*(x^o,u^o,t)\psi(t) + \Phi_u(x^o,u^o,t), \bar{u}(t)\rangle\, dt = 0\ \forall \bar{u}.$$

Logo

$$\langle -\varphi_u^*(x^o,u^o,t)\psi(t) + \Phi_u(x^o,u^o,t), \bar{u}\rangle = 0 \text{ q.t.p. } 0 \le t \le T.$$

Desta forma a desigualdade (4.8) é satisfeita.

2) Se o sistema (4.10) for degenerado, existiria uma função $\psi(t)$ não-nula solução da equação (4.7) com $\lambda_0 = 0$ tal que $-\varphi_u^*(x^o,u^o,t)\psi(t) \equiv 0$, que recai na desigualdade (4.8) com $\lambda_0 = 0$. □

Observação 4.2. *Introduza a função*

$$H(x,u,\psi,t) = \langle \varphi(x,u,t), \psi\rangle - \lambda_0\Phi(x,u,t).$$

Chamada de função de Pontryagin ou também Hamiltoniano. Então,

$$H_u(x^o,u^o,\psi,t) = \langle \varphi_u^*(x^o,u^o,t), \psi(t)\rangle - \lambda_0\Phi_u(x^o,u^o,t),$$

e como uma condição necessária para $H(x^o,u,\psi,t)$ atingir seu máximo em M como função

de u, é que $-H_u(x^o, u^o, \psi, t)$ seja suporte para M em $u^\circ(t)$. Logo a condição (4.8) pode ser reescrita como segue.
Se (x^o, u^o) é uma solução do Problema 4.1, então $H(x^o, u, \psi, t)$ como função de u em M satisfaz as condições necessárias para um máximo no ponto $u = u^o(t)$ q.t.p. $0 \leq t \leq T$.

Nas observações a seguir discutimos algumas modificações nas condições de contorno para diferentes versões de problemas de controle ótimo:

Observação 4.3. *Substituindo a condição de contorno (4.3) por uma restrição mais geral*

$$G_i(x(T)) = 0, \ i = 1, ..., k \tag{4.16}$$

onde G_i são funções escalares diferenciáveis em $\mathbb{R}^n$. Depois, utilizando os mesmos argumento como no Exemplo 3.4 e assumindo que verifica-se a condição de não degeneração, o subespaço tangente $T(Q_2, (x^o, u^o))$ ao conjunto Q_2 no ponto (x^o, u^o) consiste de todos os pares $(\bar{x}, \bar{u}) \in E$, tais que:

$$\begin{aligned} &\frac{d\bar{x}(t)}{dt} = \varphi_x(x^o, u^o, t)\bar{x} + \varphi_u(x^o, u^o, t)\bar{u}, \ \bar{x}(0) = 0, \\ &\langle G_i'(x^o(T)), \bar{x}(T)\rangle = 0, \ i = 1, ..., k. \end{aligned} \tag{4.17}$$

Neste caso o operador Λ é dado por

$$\Lambda(x, u) = \left(x(t) - c - \int_0^t \varphi(x(\tau), u(\tau), \tau) d\tau, G_1(x(T)), \ldots, G_k(x(T)) \right).$$

Agora assuma que $\psi(t)$ satisfaça a equação diferencial (4.7) com a condições de fronteira

$$\psi(T) = \sum_{i=1}^{k} \lambda_i G_i'(x^o(T)), \tag{4.18}$$

onde λ_i são números arbitrários. Assim podemos utilizar exatamente os mesmos argumentos para provar (4.8). Assim, a formulação das condições necessárias de extremo serão da forma (4.8), com a única diferença de que a função $\psi(t)$ que aparece aí tem que satisfazer também as condições de contorno (4.18).

Observação 4.4. *Em geral, se o estado $x(t)$ é sujeito a condições de contorno na forma general $x(0) \in S_1$, $x(T) \in S_2$, onde S_1 e S_2 são variedades diferenciáveis, então a condição (4.8) pode ser mantida, mas com hipóteses adicionais para $\psi(0)$ e $\psi(T)$: $\psi(0)$ tem que ser transversal a S_1 no ponto $x^o(0)$, e $\psi(T)$ transversal a S_2 em $x^o(T)$.*
Em particular, se o extremo esquerdo é fixado, $x(0) = c$, enquanto o extremo direito é livre, então as condições de contorno precisam ser impostas só em $\psi(T)$, em virtude do fato de que, durante a prova não teríamos o conjunto L_2 e por isto na equação (4.12)

não apareceria o termo $f_2(\bar{x}, \bar{u})$*. Assim continuando com o mesmo raciocínio da demonstração, temos que impor* $\psi(T) = 0$*, por causa do termo* $\langle \psi(t), \bar{x}(t) \rangle|_0^T$*. Uma importante observação aqui é que neste caso especial a condição de não degeneração é supérfluo, então podemos assumir* $\lambda_0 = 1$*.*

Modificações análogas das condições extremais verificam-se para o caso em que S_1 *e* S_2 *são conjuntos convexos.* $\psi(0)$ *tem que ser o suporte de* S_1 *no ponto* $x^o(0)$ *e* $\psi(T)$ *o suporte de* S_2 *em* $x^o(T)$*.*

Observação 4.5. *Se o funcional a ser minimizado não é um funcional integral do tipo (4.4), mas é do tipo* $\Phi_0(x(T))$*, ou seja, é um problema tipo Mayer, onde* $\Phi_0(x)$ *é uma função diferenciável em* $\mathbb{R}^n$*, então este problema se reduz ao problema 4.1, desde que*

$$\Phi_0(x(T)) = \Phi_0(x(0)) + \int_0^T \langle \Phi_0'(x(t)), \varphi(x(t), u(t), t) \rangle dt.$$

Entretanto, problemas deste tipo podem ser considerados diretamente, desde que o cone de direções de descida para este funcional é determinado de forma muito simples. Verifica-se que as condições de extremo tem a mesma forma; basta tomar $\Phi(x, u, t) = 0$ *e* $\psi(x(T)) = -\Phi_0'(x^0(T))$*.*

CAPÍTULO 5

PRINCÍPIO DO MÁXIMO DE PONTRYAGIN

Os resultados do capítulo anterior dão condições necessárias para um mínimo local no espaço L_∞ (com u variável). No cálculo das variações tais condições são chamados de “condições de estremo fraco”(eles são obtidos comparando uma trajetória ótima com outras trajetórias numa vizinhança fraca, i.e., trajetórias que se aproximam ao ótimo uniformemente em x e u.)
Em muitos casos é mais natural considerar o problema em espaços mais gerais. Por exemplo, se $M \subset \mathbb{R}^r$ (condição (4.5) $u(t) \in M$ q.t.p. $0 \le t \le T$) consiste em exatamente de 2 pontos, e $u^\circ(t)$ é uma função arbitraria que também satisfaz (4.5), então não existe trajetória alguma perto de $u^\circ(t)$ (na norma de L_∞) que também satisfaz (4.5). Por outro lado, se o controle $u^\circ(t)$ recebe um incremento $\bar{u}(t)$ que não é zero sobre um intervalo pequeno de tempo Δt, que denominaremos “Spike Variation”(veja a Figura 5.1), então o incremento para o espaço de fase também será pequena (mesmo quando $\|\bar{u}\|_{L_\infty}$ não seja pequeno). Portanto é desejável considerar o problema num espaço para o qual o tamanho do “Spike Variation”seja pequena (por exemplo o espaço L_1). As condições para extremo local nesse tipo de espaços são chamados de “condições de extremo forte”.

Neste capítulo consideramos a derivação das condições para este tipo de Problemas de Controle. Ao fazer isso é possível relaxar a condição de que M seja convexo e conter pontos interiores, nem precisamos assumir que $\Phi(x, u, t)$ e $\varphi(x, u, t)$ sejam diferenciáveis com respeito a u.
Em vez de analisar diretamente o funcional e as restrições no espaço L_1, empregaremos um “método artificial”, sugerido por Dubovitskii- Miltyutin. A ideia principal deste método é introduzir uma transformação na variável temporal. Isto será feito de tal forma que uma

pequena variação da transformação corresponde a uma variação do "Spike Variation"no problema inicial.

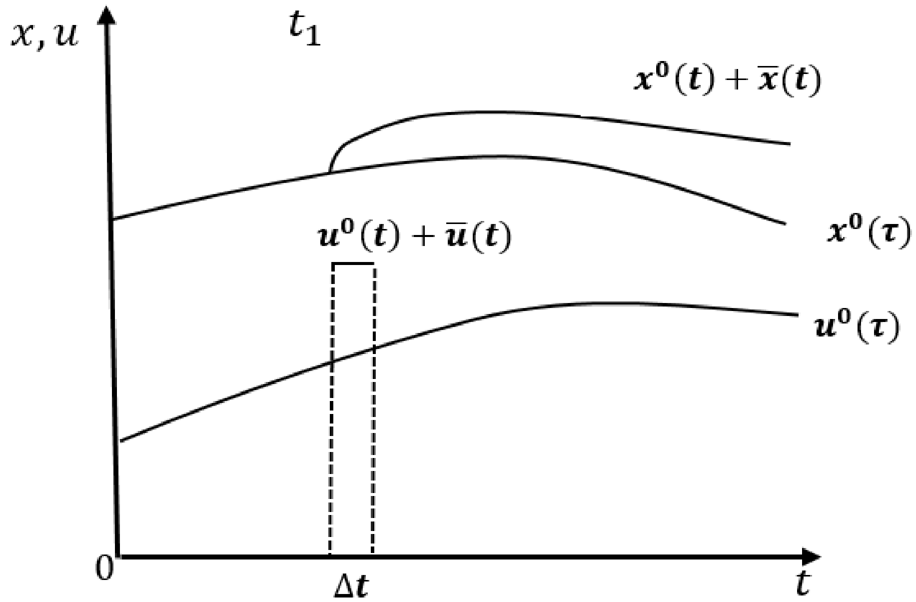


Figura 5.1: Spike Variation

Vamos proceder com a formulação rigorosa do problema.

5.1 Princípio de Máximo para problemas autônomos com tempo final livre

Problema 5.1. *Encontrar funções* $x : [t_0, t_1] \to \mathbb{R}^n$, *(trajetória de fase) e* $u : [0, T] \to \mathbb{R}^r$ *(controle) que satisfaçam certas condições, ou seja,*

$$\min F(x, u) = \int_{t_0}^{t_1} \Phi(x(t), u(t))dt \tag{5.1}$$

$s.a$

$$\frac{dx(t)}{dt} = \varphi(x(t), u(t)) \ \ q.t.p. \ \ t_0 \leq t \leq t_1, \tag{5.2}$$

$$x(t_0) = c \ \ e \ \ x(t_1) = d, \tag{5.3}$$

$$u(t) \in M \ \ q.t.p. \ \ t_0 \leq t \leq t_1. \tag{5.4}$$

O funcional é minimizado sobre todo $t_1 \geq t_0$, *todas as funções* $u(t)$ *limitadas e mensuráveis satisfazendo (5.4), todas as* $x(t)$ *continuas satisfazendo (5.2)-(5.3).*

Condições do Problema:

- $\Phi(x(t), u(t))$ e $\varphi(x(t), u(t))$ definidas em $\mathbb{R}^{n+r}$ com valores em $\mathbb{R}^n$ e $\mathbb{R}^1$ respectivamente, contínuas em u, continuamente diferenciáveis em x.
- $\Phi_x(x(t), u(t))$ e $\varphi_x(x(t), u(t))$ são limitadas para todos os $x(t)$,$u(t)$ limitadas e arbitrárias.

Diferenças entre os Problema 5.1 e 4.1:

- No Problema 4.1 consideramos M como sendo convexo, fechado e $int(M) \neq \emptyset$, enquanto no Problema 5.1 $M \subset \mathbb{R}^r$ (qualquera).
- Em 5.1 não requer que $\Phi(x(t), u(t))$, $\varphi(x(t), u(t))$ sejam diferenciáveis com respeito a u. Isto é natural tendo em conta que M pode consistir de um número finito de pontos.
- As funções $\Phi(x(t), u(t))$, $\varphi(x(t), u(t))$ não dependem explicitamente de t.
- O tempo t_1 em 5.1 não é fixo a diferença de que no Problema 4.1 T é fixo.

Observação 5.1. *O sistema descrito acima é conhecido como "sistema invariante no tempo".*

Agora introduzimos uma transformação do tempo $\tau \longmapsto t$,

$$t : [0, 1] \longrightarrow [t_0, t_1],$$

aplicação de $[0, 1]$ sobre $[t_0, t_1]$, definida por certa função $v(\tau) \in L_1([0, 1]; \mathbb{R})$ como sendo:

$$t(\tau) = t_0 + \int_0^\tau v(t)dt, \tag{5.5}$$

$$t(1) = t_1, \tag{5.6}$$

$$v(\tau) \geq 0. \tag{5.7}$$

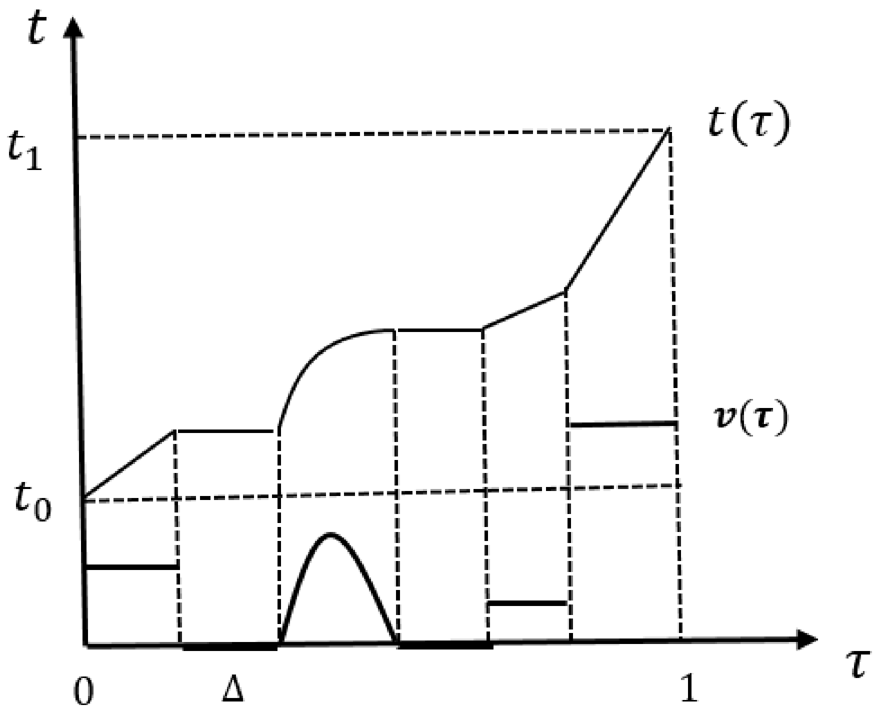


Figura 5.2: Transformação temporal $t(\tau)$

A transformação assim definida é injetiva se $v(\tau) > 0$ para todo $0 \leq \tau \leq 1$. De fato, suponha por absurdo que t não seja injetiva, então, existem $\widehat{\tau}, \widetilde{\tau} \in [0,1]$, $\widehat{\tau} \neq \widetilde{\tau}$, tais que $t(\widehat{\tau}) = t(\widetilde{\tau})$. Logo,

$$\int_0^{\widehat{\tau}} v(t)dt = \int_0^{\widetilde{\tau}} v(t)dt \Longrightarrow \int_0^{\widetilde{\tau}} v(t)dt - \int_0^{\widehat{\tau}} v(t)dt = 0.$$

Supondo $\widehat{\tau} < \widetilde{\tau}$, temos que,

$$\int_{\widehat{\tau}}^{\widetilde{\tau}} v(t)dt + \int_0^{\widetilde{\tau}} v(t)dt - \int_0^{\widetilde{\tau}} v(t)dt = 0$$

$$\Longrightarrow \int_{\widehat{\tau}}^{\widetilde{\tau}} v(t)dt = 0,$$

insto implica que $v(\tau) = 0,\ \tau \in [\widehat{\tau}, \widetilde{\tau}]$, o que contradiz a suposição de que $v(\tau) > 0$, para todo $\tau \in [0,1]$.

Agora, se $v(\tau)$ anula-se sobre algum intervalo fechado Δ, então $t(\tau) = const.$, para $\tau \in \Delta$, pois

$$t(\tau) = t_0 + \int_0^{\tau} \underbrace{v(t)}_{=0} dt = t_0.$$

Para fazer $\tau(t)$ injetora suporemos que

$$\tau(s) = \inf\{\tau : t(\tau) = s\}. \tag{5.8}$$

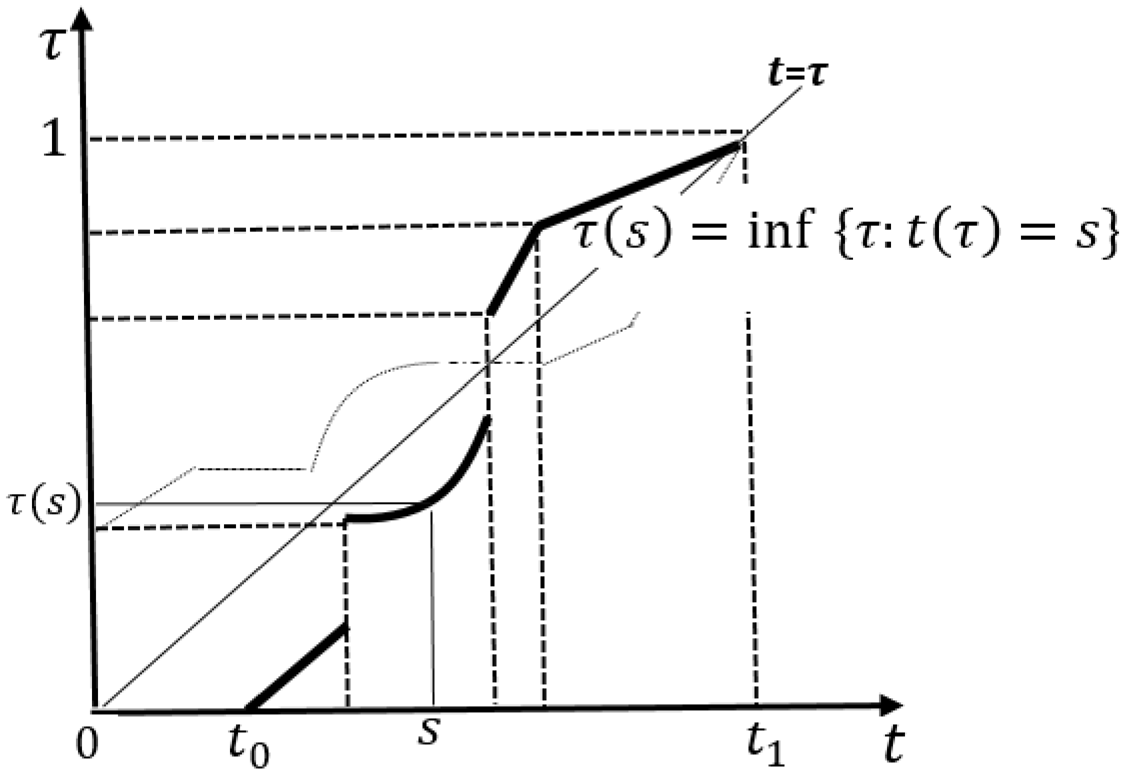


Figura 5.3: Inversa da transformação $t(\tau)$

Consideremos agora um problema de controle ótimo auxiliar nas variáveis τ, v, x para $v(\tau)$ definida (no sentido que $v(\tau)$ é dado), satisfazendo (5.5)-(5.7).

Problema 5.2.

$$\min I(x,u,v) = \int_0^1 v(\tau)\Phi(x(\tau),u(\tau))d\tau, \tag{5.9}$$

$$\textit{com } x(\tau) \in C\left([0,1];\mathbb{R}^n\right),\ u(\tau) \in L_\infty\left([0,1];\mathbb{R}^r\right),\quad v(\tau) \in L_1\left([0,1];\mathbb{R}\right),$$

$s.a$

$$\frac{dx(\tau)}{d\tau} = v(\tau)\varphi(x(\tau),u(\tau))\ \textit{q.t.p. } 0 \leq \tau \leq 1, \tag{5.10}$$

$$x(0) = c, x(1) = d, \tag{5.11}$$

$$v(\tau) \geq 0, \tag{5.12}$$

$$u(\tau) \in M,\ \textit{q.t.p. } 0 \leq \tau \leq 1. \tag{5.13}$$

Verificamos inicialmente que toda solução $(x(t),u(t))$ do Problema 5.1 induz uma solução $(x(\tau),u(\tau))$ do problema 5.2 e reciprocamente. O resultado a seguir relaciona os problemas 5.1 e 5.2.

Lema 5.1. *Sejam $x(t)$, $u(t)$ funções satisfazendo (5.2) e as restrições (5.3), (5.4). Então para qualquer $v(\tau)$ satisfazendo (5.5)-(5.7), as funções*

$$x(\tau) = x(t(\tau)),$$

$$u(\tau) = \begin{cases} u(t(\tau)), & \textit{para } \tau \in R_1, \\ \textit{arbitrário}, & \textit{para } \tau \in R_2, \end{cases} \tag{5.14}$$

onde, $t(\tau)$ é definido por (5.5) e

$$R_1 = \{\tau \in [0,1] : v(\tau) > 0\},\ R_2 = \{\tau \in [0,1] : v(\tau) = 0\},$$

satisfazem a equação (5.10) e as restrições (5.11)-(5.13). Além disso,

$$I(x,u,v) = F(x,u).$$

Reciprocamente, se $x(\tau)$,$u(\tau)$,$v(\tau)$ satisfazem (5.10)-(5.13), então as funções definidas por:

$$x(t) = x(\tau(t)),$$

$$u(t) = u(\tau(t)),$$

com

$$t_1 = t_0 + \int_0^1 v(\tau)d\tau$$

e $\tau(t)$ é dado por (5.8), satisfazem (5.2)-(5.4). Além disso,

$$F(x,u) = I(x,u,v).$$

Demonstração. Supondo que $x(t)$, $u(t)$ satisfazem (5.2) e as restrições (5.3), (5.4) e seja $v(\tau)$ satisfazendo (5.5)-(5.7), então:

i) Se $\tau \in R_1$,

$$\begin{aligned} \frac{dx(\tau)}{d\tau} = \frac{dx(t(\tau))}{d\tau} &= \frac{dx(t(\tau))}{dt}\frac{dt(\tau)}{d\tau} \\ &= v(\tau)\varphi(x(t(\tau)), u(t(\tau))) \\ &= v(\tau)\varphi(x(\tau), u(\tau)). \end{aligned}$$

$$\Longrightarrow \frac{dx(\tau)}{d\tau} = v(\tau)\varphi(x(\tau), u(\tau)).$$

ii) Se $\tau \in R_2$, temos que $v(\tau) = 0$ e $t(\tau) = const.$, de modo que $x(\tau) = x(t(\tau)) = const.$

$$\Longrightarrow \frac{dx(\tau)}{d\tau} = 0 = v(\tau)\varphi(x(\tau), u(\tau)).$$

Agora mostremos a igualdade dos funcionais. Para isto, seja $\Psi(t) := \Phi(x(t), u(t))$, $t \in [t_0, t_1]$.

Observe que $\Phi(\cdot, u(t))$ é contínua, portanto mensurável, além disso, $u(t)$ é mensurável, portanto $\Phi(x(t), \cdot)$ é mensurável, donde $\Psi(t)$ é mensurável. Logo usando o teorema de mudança de variáveis para integrais de Lebesque (ver apêndice D) temos que:

$$\begin{aligned} F(x,u) &= \int_{t_0}^{t_1} \Psi(t)dt \\ &= \int_0^1 \Psi(t(\tau))t'(\tau)d\tau \\ &= \int_0^1 \Phi(x(t(\tau)), u(t(\tau)))v(\tau)d\tau \\ &= \int_0^1 v(\tau)\Phi(x(\tau), u(\tau))d\tau \\ &= I(x,u,v). \end{aligned}$$

Sejam $x(\tau), u(\tau), v(\tau)$ tais que satisfazem (5.7)-(5.3) e $x(t) = x(\tau(t))$, $u(t) = u(\tau(t))$ e seja $t \in [t_0, t_1]$ tal que $\tau(t) \in R_2$. Então

$$v(\tau(t)) = 0 \text{ e } \frac{d}{d\tau}x(\tau(t)) = v(\tau(t))\varphi(x(\tau(t)), u(\tau(t))) = 0,$$

portanto, $x(t) = const.$ e $x(\tau(t)) = const.$
Por outro lado, considerando $t \in [t_0, t_1]$ com $\tau(t) \in R_1$ temos que

$$\begin{aligned}
\frac{dx(t)}{dt} &= \frac{dx(\tau(t))}{dt} \\
&= \frac{dx(\tau(t))}{d\tau}\frac{d\tau(t)}{dt} \\
&= v(\tau)\varphi(x(\tau(t)), u(\tau(t)))\frac{1}{v(\tau(t))} \\
&= \varphi(x(t), u(t)).
\end{aligned}$$

Isto mostra que a equação (5.10) se transforma em (5.2).
Para a igualdade dos funcionais, note que

$$v(\tau)\varphi(x(\tau), u(\tau)) = 0,$$

para todo $\tau \in R_2$. Além disso, $\mu\left(\tau^{-1}(R_2) \cap [t_0, t_1]\right) = 0$, onde μ é a medida de Lebesque em $\mathbb{R}$, então

$$\int_{\tau^{-1}(R_2)\cap[t_0,t_1]} \varphi(x(t), u(t), t)dt = 0.$$

Logo,

$$\begin{aligned}
I(x, u, v) &= \int_0^1 v(\tau)\varphi(x(\tau), u(\tau))d\tau \\
&= \int_{R_1} v(\tau)\varphi(x(\tau), u(\tau))d\tau + \int_{R_2} v(\tau)\varphi(x(\tau), u(\tau))d\tau \\
&= \int_{\tau^{-1}(R_1)\cap[t_0,t_1]} v(\tau(t))\varphi(x(\tau(t)), u(\tau(t)))d\tau(t) + 0 \\
&= \int_{\tau^{-1}(R_1)\cap[t_0,t_1]} v(\tau(t))\varphi(x(t), u(t), t)\frac{1}{v(\tau(t))}dt + 0 \\
&= \int_{\tau^{-1}(R_1)\cap[t_0,t_1]} \varphi(x(t), u(t), t)dt + \int_{\tau^{-1}(R_2)\cap[t_0,t_1]} \varphi(x(t), u(t), t)dt \\
&= \int_{t_0}^{t_1} \varphi(x(t), u(t), t)dt \\
&= F(x, u).
\end{aligned}$$

Assim, fica provado o Lema. □

Corolário 5.1. *Se $x^o(t), u^o(t)$ é solução do Problema 5.1, então para qualquer $v^o(\tau)$ satisfazendo (5.5)-(5.7) as funções $x^o(\tau), u^o(\tau), v^o(\tau)$ definidas por (5.14) fornecem uma solução do Problema 5.2. Naturalmente, este também fornece uma solução do novo Problema 5.3, que é igual ao Problema 5.2 exceto que $u^o(\tau)$ é fixo e $I(x, u^o, v)$ é minimizado só sobre $x(\tau)$ e $v(\tau)$.*

Para estabelecer um Princípio de Máximo para o Problema 5.1 usamos a transformação do tempo 5.5, logo através da escolha adequada de u^o e v^o obtemos a condição do máximo para o Problema 5.2 que é fornecido pelo Teorema 4.1, e finalmente usando o Lema 5.1 obtemos o resultado desejado. A dedução do Princípio do Máximo para o Problema 5.1 esta dividido em varias etapas:

Consideremos $x^o(t)$, $u^o(t)$, t_1 como sendo solução do problema inicial (5.1) e seja $v^o(\tau)$ qualquer função satisfazendo (5.6) e (5.7). A forma dessa função será definida mais adiante.
Definimos $u^o(\tau)$ conforme (5.14), isto é:

$$u^o(\tau) = \begin{cases} u^o(t(\tau)), & \text{para } \tau \in R_1, \\ \text{arbitrário}, & \text{para } \tau \in R_2. \end{cases}$$

A definição de $u^o(\tau)$ quando $\tau \in R_2$ será dada mais adiante. Então vamos resolver o novo problema:

Problema 5.3.

$$\min_{x,v} F(x, u^o, v) = \int_0^1 v(\tau)\Phi(x(\tau), u^o(\tau))d\tau \tag{5.15}$$

$$\textit{sujeita a: } \frac{dx(\tau)}{d\tau} = v(\tau)\varphi(x(\tau), u^o(\tau)), \ \ q.t.p. \ 0 \leq \tau \leq 1, \tag{5.16}$$

$$x(1) = d, \tag{5.17}$$

$$v(\tau) \geq 0, \ \ \tau \in [0, 1]. \tag{5.18}$$

Pelo Corolário 5.1, $x^o(\tau)$ e $v^o(\tau)$ constituem uma solução do Problema 5.3. Para derivar uma condição necessária de otimalidade para este problema, usaremos o resultado do capítulo anterior (Teorema 4.1), com $v(t)$ como sendo o controle. De fato este resultado é aplicável, pois:

i) A restrição (5.18) é de forma $v(\tau) \in M_1$, $0 \leq \tau \leq 1$, onde $M_1 = [0, +\infty[$ é um conjunto convexo em $\mathbb{R}$ com interior não vazio (semi-eixo positivo);

ii) O integrando em (5.15) e a parte direita de (5.16) são diferenciáveis com respeito a v;

iii) O tempo $T = 1$ é fixo.

Portanto, do Teorema 4.1 (teorema do princípio do máximo local para o Problema

5.3), existem $\psi(\tau)$ e $\lambda_0 \geq 0$ não simultaneamente nulos tais que,

$$\frac{d\psi(\tau)}{d\tau} = -v^o(\tau)\varphi_x^*(x^o(\tau), u^o(\tau))\psi(\tau) + \lambda_0 v^o(\tau)\Phi_x(x^o(\tau), u^o(\tau)), \tag{5.19}$$

$$[-\langle \varphi(x^o(\tau), u^o(\tau)), \psi(\tau)\rangle + \lambda_0 \Phi(x^o(\tau), u^o(\tau)](v - v^o(\tau)) \geq 0 \tag{5.20}$$

para quase todos $0 \leq \tau \leq 1$ e para todo $v(\tau) \geq 0$.
Da ultima desigualdade segue que

$$-\langle \varphi(x^o(\tau), u^o(\tau)), \psi(\tau)\rangle + \lambda_0 \Phi(x^o(\tau), u^o(\tau)) = 0, \text{ q.t.p. } \tau \in R_1 \tag{5.21}$$

$$-\langle \varphi(x^o(\tau), u^o(\tau)), \psi(\tau)\rangle + \lambda_0 \Phi(x^o(\tau), u^o(\tau)) \geq 0, \text{ q.t.p. } \tau \in R_2 \tag{5.22}$$

De fato, seja $\tau \in R_1$, tal que vale (5.20) e seja

$$\alpha = -\langle \varphi(x^o(\tau), u^o(\tau)), \psi(\tau)\rangle + \lambda_0 \Phi(x^o(\tau), u^o(\tau)),$$

então (5.20) torna-se

$$\alpha(v - v^o(\tau)) \geq 0, \text{para todo } v \geq 0.$$

Suponhamos que a igualdade (5.21) é falsa, i.e., $\alpha \neq 0$, então $\alpha > 0$ ou $\alpha < 0$.

i) Se $\alpha > 0$, e como a desigualdade (5.20) vale para todo $v \geq 0$, podemos escolher v de tal forma que $v - v^o(\tau) < 0$, logo teríamos que $\alpha(v - v^o(\tau)) < 0$ o que é um absurdo.

ii) Da mesma forma, se $\alpha < 0$ escolhemos v de tal forma que $v - v^o(\tau) > 0$, logo temos que $\alpha(v - v^o(\tau)) < 0$, o que é um absurdo outra vez.

Portanto, $\alpha = -\langle \varphi(x^o(\tau), u^o(\tau)), \psi(\tau)\rangle + \lambda_0 \Phi(x^o(\tau), u^o(\tau) = 0$ para $\tau \in R_1$ $q.t.p.$ Assim a igualdade (5.21) é verificada. A desigualdade (5.22) é válida desde que $v^o(\tau) = 0$ quando $\tau \in R_2$.
Aqui, $R_1 = \{\tau \in [0, 1] : v^o(\tau) > 0\}$ e $R_2 = \{\tau \in [0, 1] : v^o(\tau) = 0\}$.

Agora introduzimos a função $\psi(t) = \psi(\tau(t))$, desde que

$$\psi(t) = const. = \psi(\tau(t))$$

quando $v^o(\tau) = 0$, segue deste fato e de (5.19) que $\psi(t)$ satisfaz

$$\frac{d\psi(t)}{dt} = -\varphi_x^*(x^o(t), u^o(t))\psi(t) + \lambda_0 \Phi_x(x^o(t), u^o(t)), \text{ q.t.p. } t_0 \leq t \leq t_1 \tag{5.23}$$

Se substituirmos t por τ na condição (5.21) temos,

$$-\langle \varphi(x^o(t), u^o(t)), \psi(t)\rangle + \lambda_0 \Phi(x^o(t), u^o(t)) = 0 \tag{5.24}$$

para quase todo ponto $t_0 \leq t \leq t_1$ tal que $t = t(\tau)$ e $\tau \in R_1$. Mas o conjunto de pontos t tal que $t = t(\tau)$ e $\tau \in R_2$ é no máximo enumerável (pois a função monótona

$$t(\tau) = t_0 + \int_0^\tau v^o(s)ds$$

é injetora em todo ponto exceto talvez sobre um conjunto enumerável). Portanto, vale a condição (5.24) para quase todo ponto $t_0 \leq t \leq t_1$.

Falta analisar a condição (5.22), mas primeiro vamos especificar a forma das funções $v^o(\tau)$ e $u^o(\tau)$ sobre R_2. Consideremos $R_1 \subset [0, 1]$ com as seguintes propriedades:

i) um subconjunto perfeito nunca denso em $[0, 1]$ (isto é, não contem intervalos, obtém-se suprimindo de [0,1] uma quantidade enumerável de intervalos);

ii) se a interseção de R_1 com qualquer intervalo é não vazia, então tem medida positiva;

iii) a construção de tal conjunto é análoga à construção do conjunto de Cantor, exceto que os intervalos devem ser suprimidas de tal forma que a medida do conjunto restante é sempre maior do que uma constante positiva. Veja Apêndice A.

Definimos $v^o(\tau)$ como:

$$v^\circ(\tau) = \begin{cases} \dfrac{t_1 - t_0}{\mu(R_1)} & \text{se } \tau \in R_1, \\ 0 & \text{se } \tau \in R_2 = [0, 1] \setminus R_1. \end{cases}$$

Logo,

$$t(1) = t_0 + \int_0^1 v^o(s)ds = t_0 + \int_{R_1} v^o(s)ds + \int_{R_2} v^o(s)ds = t_0 + \frac{t_1 - t_0}{\mu(R_1)} \int_{R_1} ds = t_1$$

e $v^o(\tau) \geq 0$.
Donde, (5.6) e (5.7) são verificados.

Agora, vamos definir $u^\circ(\tau)$ para $\tau \in R_2$. O conjunto R_2 como complemento do conjunto perfeito R_1 é a união enumerável de um conjunto de intervalos. Seja $\tau \in \Delta$, onde Δ é um desses intervalos. É possível escrever Δ como união disjunta enumerável de intervalos fechados à esquerda:

$$\Delta = \cup_{i=1}^\infty \Delta_i.$$

Tome $u^i \in M$, $i \in \mathbb{N}$, onde $\{u^i \in M : i \in \mathbb{N}\}$ constitui um subconjunto enumerável e denso em M (isto é sempre possível, desde que $\mathbb{R}^r$ é separável). Defina $u^\circ(\tau) = u^i, \tau \in \Delta_i$. Isto completa a definição de $v^\circ(\tau)$ e $u^\circ(\tau)$.

Finalmente, analisemos a condição (5.22). Para isso escrevemos esta desigualdade para $\tau \in \Delta_i \subseteq \Delta \subseteq R_2$, com a mesma notação de antes:

$$-\left\langle \varphi(x^\circ(\tau), u^i)), \psi(\tau) \right\rangle + \lambda_0 \Phi(x^\circ(\tau), u^i) \geq 0 \;\; \tau \in \Delta_i.$$

Mas,

$$x^\circ(\tau) = const. = x^\circ(\tau(t)), \tau \in \Delta \text{ e } \psi(\tau) = const. = \psi(\tau(t)), \tau \in \Delta.$$

Assim,

$$-\left\langle \varphi(x^\circ(\tau(t)), u^i)), \psi(\tau(t)) \right\rangle + \lambda_0 \Phi(x^\circ(\tau(t)), u^i) \geq 0, \;\; i \in \mathbb{N}.$$

Como $\{u^i : i \in \mathbb{N}\}$ é denso em M e

$$u \longmapsto \langle \varphi(x^\circ(\tau(t)), u), \psi(\tau(t)) \rangle + \lambda_0 \Phi(x^\circ(\tau(t)), u)$$

é contínua, segue que

$$-\langle \varphi(x^\circ(\tau(t)), u), \psi(\tau(t)) \rangle + \lambda_0 \Phi(x^\circ(\tau(t)), u) \geq 0, \tag{5.25}$$

para todo $u \in M$ e para quase todo $\tau \in R_2$.
Tomando $x^\circ(t) = x^\circ(\tau(t))$ e $\psi(t) = \psi(\tau(t))$ temos que,

$$-\langle \varphi(x^\circ(t), u), \psi(t) \rangle + \lambda_0 \Phi(x^\circ(t), u) \geq 0 \;\; \forall u \in M. \tag{5.26}$$

Esta desigualdade é válida para todo t tal que a medida do conjunto $R_t = \{\tau \in R_2 : t(\tau) = t\}$ é positiva. De fato, se $\mu(R_t) > 0$, então a desigualdade (5.25) é válido para algum $\tau \in R_t$(já que (5.25) deixa de ser verdade só num conjunto de medida nula).
O conjunto de pontos $t \in [t_0, t_1]$ tais que $\mu(R_t) > 0$ é denso em $[t_0, t_1]$
De fato, se este contem algum segmento, então sua pre imagem sob $t(\tau)$ é também um segmento, mas R_2 é denso, e a interseção com esse segmento é não vazio e contem um segmento Δ(R_2 união enumerável de intervalos). Seja $\tau \in \Delta$, $t = t(\tau)$; então $R_t \supset \Delta$, assim $\mu(R_\tau) > 0$, contradição.
Portanto o conjunto de todos os t tal que vale (5.26) é denso em $[t_0, t_1]$. Além disso,

$$t \longmapsto \langle \varphi(x^\circ(t), u), \psi(t) \rangle + \lambda_0 \Phi(x^\circ(t), u)$$

é contínua como função de t. Portanto (5.26) é válido para todo $[t_0, t_1]$.

Desta forma mostramos o seguinte resultado:

Teorema 5.1. *(Princípio do Máximo de Pontryagin.) Suponha que as condições (5.1)-(5.4) do Problema 5.1 sejam válidas. Seja $x^o(t)$, $u^o(t)$, t_1 uma solução do Problema (5.1).*

Então existem $\psi(t)$ e λ_0 ambos não simultaneamente nulos, tais que:

$$\frac{d\psi(t)}{dt} = -\varphi_x^*(x^o(t), u^o(t))\psi + \lambda_0 \Phi_x(x^o(t), u^o(t)) \ q.t.p. \ t_0 \leq t \leq t_1, \tag{5.27}$$

$$H(x^o(t), u^o(t), \psi(t), t) = 0 \ q.t.p. \ t_0 \leq t \leq t_1, \tag{5.28}$$

$$H(x^o(t), u(t), \psi(t), t) \leq 0 \ \forall u \in M, \forall t \in [t_0, t_1], \tag{5.29}$$

onde

$$H(x, u, \psi, t) = \langle \varphi(x, u), \psi(t) \rangle - \lambda_0 \Phi(x, u). \tag{5.30}$$

Em outras palavras, a função $H(x^o, u, \psi(t), t)$ assume o máximo sobre $u \in M$ em $u = u^o(t)$ em quase todo ponto de $t_0 \leq t \leq t_1$.

5.2 Princípio de Máximo para problemas não autônomos com tempo final livre

Nesta seção assumimos que as funções Φ e φ do Problema 5.1 dependem explicitamente de t, isto é, $\Phi = \Phi(x, u, t)$, $\varphi = \psi(x, u, t)$. Além disso, $\Phi(x, u, t)$ e $\varphi(x, u, t)$ satisfazem as condiçoes(5.1-(5.2 e são continuamente diferenciáveis com respeito a t. Então estamos no seguinte problema de controle.

Problema 5.4.

$$\begin{cases} \min F(x, u) = \displaystyle\int_{t_0}^{t_1} \Phi(x(t), u(t), t) dt \\ s.a \\ \qquad \dfrac{dx(t)}{dt} = \varphi(x(t), u(t), t) \ q.t.p. \ t_0 \leq t \leq t_1, \\ \qquad x(t_0) = c, \ x(t_1) = d, \\ \qquad u(t) \in M. \end{cases}$$

Este problema será reduzido ao Problema 5.1. Para isso completamos às variáveis $x = (x_1(t), x_2(t), \ldots, x_n(t))$ com uma nova variável $x_{n+1}(t)$, satisfazendo

$$\frac{dx_{n+1}(t)}{dt} = 1, \ x_{n+1}(t_0) = t_0.$$

Então o Problema 5.4 pode ser escrita como:

$$\begin{cases} \min F(x,u) = \displaystyle\int_{t_0}^{t_1} \Phi(x(t),u(t),x_{n+1}(t))dt \\ s.a \\ \qquad \dfrac{dx(t)}{dt} = \varphi(x(t),u(t),x_{n+1}(t)), \text{ q.t.p.} t_0 \le t \le t_1, \\ \qquad x(t_0) = c, \\ \qquad \dfrac{dx_{n+1}(t)}{dt} = 1 \text{ q.t.p. } t_0 \le t \le t_1, \; x_{n+1}(t_0) = t_0, \\ \qquad u(t) \in M. \end{cases}$$

Definimos,

$$\begin{aligned} \widetilde{x}(t) &:= (x(t), x_{n+1}(t)), \\ \widetilde{\Phi}(\widetilde{x},u) &:= \Phi(x(t),u(t),x_{n+1}(t)), \\ \widetilde{\varphi}(\widetilde{x},u) &:= (\varphi(x(t),u(t),x_{n+1}(t)),1)\,. \end{aligned}$$

Mais ainda, o problema anterior pode ser escrita como

$$(PA)\begin{cases} \min \widetilde{F}(x,u) = \displaystyle\int_{t_0}^{t_1} \widetilde{\Phi}(\widetilde{x}(t),u(t))dt \\ s.a \\ \qquad \dfrac{d\widetilde{x}(t)}{dt} = \widetilde{\varphi}(\widetilde{x}(t),u(t)), \\ \qquad \widetilde{x}(t_0) = (c,t_0), \widetilde{x}(t_1) = (d,0), \\ \qquad u(t) \in M. \end{cases}$$

Note que resolver o Problema 5.4 é equivalente a resolver o Problema (PA), e este último satisfaz todas as hipóteses do princípio de Máximo, Teorema 5.1, ou seja, $\widetilde{\Phi}(\widetilde{x},u(t))$ e $\widetilde{\varphi}(\widetilde{x},u(t))$ são contínuas em $\widetilde{x}$ e u, continuamente diferenciáveis com respeito a $\widetilde{x}$, além disso $\widetilde{\Phi}_{\widetilde{x}}(\widetilde{x}(t),u(t))$ e $\widetilde{\varphi}_u(\widetilde{x}(t),u(t))$ são limitados. Aplicando o Teorema 5.1 ao Problema (PA) segue que:

Se $x^o(t)$, $u^o(t)$, $x^o_{n+1}(t)$ é uma solução do Problema (PA), existem $\widetilde{\psi}(t) = (\psi_1(t),\psi_2(t),\ldots,\psi_{n+1}(t))$ e $\lambda_0 \ge 0$ tais que,

$$\frac{d\widetilde{\psi}(t)}{dt} = -\widetilde{\varphi}^*_{\widetilde{x}}(\widetilde{x}^o(t),u^o(t))\widetilde{\psi} + \lambda_0\widetilde{\Phi}_{\widetilde{x}}(\widetilde{x}^o(t),u^o(t)) \text{ q.t.p. } t_0 \le t \le t_1, \tag{5.31}$$

$$\begin{aligned} &\tfrac{d\widetilde{\psi}(t)}{dt} = -\widetilde{\varphi}^*_{\widetilde{x}}(\widetilde{x}^o(t),u^o(t))\widetilde{\psi} + \lambda_0\widetilde{\Phi}_{\widetilde{x}}(\widetilde{x}^o(t),u^o(t)) \text{ q.t.p. } t_0 \le t \le t_1, \\ &\widetilde{H}(\widetilde{x}^o(t),u^o(t),\widetilde{\psi}(t),t) = 0 \text{ q.t.p. } t_0 \le t \le t_1, \\ &\widetilde{H}(\widetilde{x}^o(t),u(t),\widetilde{\psi}(t),t) \le 0; \; \forall u \in M, \forall t \in [t_0,t_1]. \end{aligned}$$

Onde

$$\widetilde{H}(\widetilde{x},u,\psi,t)=\left\langle \widetilde{\varphi}(\widetilde{x},u),\widetilde{\psi}\right\rangle-\lambda_0\widetilde{\Phi}(\widetilde{x},u)$$

e

$$\widetilde{\varphi}_{\widetilde{x}}(\widetilde{x}^o,u^o,t)=\begin{bmatrix}\varphi_x(x^o,u^o,t) & \varphi_t(x^o,u^o,t)\\ 0 & 0\end{bmatrix},\ \widetilde{\Phi}_{\widetilde{x}}(\widetilde{x}^o,u^o,t)=\begin{bmatrix}\Phi_x(x^o,u^o,t)\\ \Phi_t(x^o,u^o,t)\end{bmatrix}.$$

De (5.31) e das equações acima segue que:

$$\frac{d\psi(t)}{dt}=-\varphi_x^*(x(t),u^o(t),t)\psi(t)+\lambda_0\Phi_x(x^o(t),u^o(t),t),$$

e

$$\frac{d\psi_{n+1}(t)}{dt}=-\left\langle\varphi_t(x^o(t),u^o(t),x^o_{n+1}(t)),\psi(t)\right\rangle+\lambda_0\Phi_t(x^o(t),u^o(t),x^o_{n+1}(t)); \tag{5.32}$$

com $\psi_{n+1}(t_1)=0$ (a condição de contorno para $\psi_{n+1}(t_1)$ é dado, desde que $x_{n+1}(t_1)$ é livre).

Além disso,

$$\begin{aligned}\widetilde{H}(\widetilde{x},u,\widetilde{\psi},t) &= \langle(\varphi(x(t),u(t),t),1),(\psi(t),\psi_{n+1}(t))\rangle-\lambda_0\Phi(x(t),u(t),t)\\ &= \langle\varphi(x(t),u(t),t),\psi(t)\rangle+\psi_{n+1}(t)-\lambda_0\Phi(x(t),u(t),t)\\ &= H(x(t),u(t),\psi(t),t)+\psi_{n+1}(t),\end{aligned}$$

e o princípio do máximo implica:

$$\begin{aligned}0 &= \max_u \widetilde{H}(\widetilde{x}(T),u,\widetilde{\psi}(t),t)\\ &= \max_u[H(x^o(t),u,\psi(t),t)+\psi_{n+1}(t)]\\ &= H(x^o(t),u^o(t),\psi(t),t)+\psi_{n+1}(t).\end{aligned}$$

$$\Rightarrow H(x^o(t),u^o(t),\psi(t),t)=-\psi_{n+1}(t)\ \text{q.t.p.}\ t_0\le t\le t_1.$$

Logo, como o lado direito da equação (5.32) é independente de ψ_{n+1}, esta equação pode ser integrada:

$$\begin{aligned}\psi_{n+1}(t_1)-\psi_{n+1}(t) &= \int_t^{t_1}\frac{d\psi_{n+1}(s)}{ds}ds\\ &= \int_t^{t_1}[-\left\langle\varphi_t(x^o(t),u^o(t),x^o_{n+1}(t)),\psi(t)\right\rangle+\lambda_0\Phi_t(x^o(t),u^o(t),x^o_{n+1}(t))]dt.\end{aligned}$$

Assim,

$$-\psi_{n+1}(t)=\int_t^{t_1}[-\langle\psi_t(x^o(t),u^o(t),t),\psi(t)\rangle+\lambda_0\Phi_t(x^o(t),u^o(t),t)]dt.$$

Logo,

$$H(x^o(t), u^o(t), \psi(t), t) = \int_t^{t_1} [-\langle \varphi_t(x^o(t), u^o(t), t), \psi(t)\rangle + \lambda_0 \Phi_t(x^o(t), u^o(t), t)]dt.$$

Assim, a condição de extremo para nosso problema tem a mesma forma que do problema autônomo, exceto que a condição

$$H(x^o(t), u^o(t), \psi(t), t) = 0$$

é substituído por

$$H(x^o(t), u^o(t), \psi(t), t) = \int_t^{t_1} [-\langle \varphi_t(x^o(t), u^o(t), t), \psi(t)\rangle + \lambda_0 \Phi_t(x^o(t), u^o(t), t)]dt.$$

O resultado a seguir é uma versão do Princípio de Máximo para problemas não autônomos com tempo final livre.

Teorema 5.2. *Suponha que as condições (5.1)-(5.2) do Problema 5.1 sejam válidas. Seja $x^o(t)$, $u^o(t)$, t_1 uma solução do Problema 5.4. Então existem $\psi(t)$ e λ_0 simultaneamente não nulos, tais que:*

$$\frac{d\psi(t)}{dt} = -\varphi_x^*(x^o(t), u^o(t), t)\psi(t) + \lambda_0 \Phi_x(x^o(t), u^o(t), t) \text{ q.t.p. } t_0 \le t \le t_1, \tag{5.33}$$

$$H(x^o(t), u^o(t), \psi(t), t) = \max_{u \in M} H(x^o(t), u(t), \psi(t), t) \text{ q.t.p. } t_0 \le t \le t_1, \tag{5.34}$$

$$H(x^o(t), u^o(t), \psi(t), t) = \int_t^{t_1} [-\langle \varphi_t(x^o(t), u^o(t), t), \psi(t)\rangle + \lambda_0 \Phi_t(x^o(t), u^o(t), t)]dt, \tag{5.35}$$

onde

$$H(x, u, \psi, t) = \langle \varphi(x, u, t), \psi(t)\rangle - \lambda_0 \Phi(x, u, t). \tag{5.36}$$

Em outras palavras, a função $H(x^o, u, \psi(t), t)$ assume o máximo sobre $u \in M$ em $u = u^o(t)$ em quase todo ponto de $t_0 \le t \le t_1$.

5.3 Princípio de Máximo para problemas não autônomos com tempo tempo final fixo

Nesta seção consideramos o Problema 5.4 com t_1 fixo. A variável $x_{n+1}(t)$ é introduzida como na seção 5.2 , que satisfará a condição de contorno $x_{n+1}(t_1) = t_1$. Portanto não haverá condição de contorno sob $\psi_{n+1}(t_1)$ na equação adjunta. Então principio do máximo terá a mesma forma da seção anterior, mas a condição (5.35) é da forma:

$$H(x^o(t), u^o(t), \psi(t), t) = \lambda + \int_t^{t_1} [-\langle \varphi_t(x^o, u^o, t), \psi(t) \rangle + \lambda_0 \Phi_t(x^o, u^o, t)]dt,$$

onde λ é uma constante.
Em particular, no caso do problema invariante no tempo conseguimos a condição

$$H(x^o(t), u^o(t), \psi(t), t) = \lambda$$

(a diferença da condição $H(x^o(t), u^o(t), \psi(t), t) = 0$ para o problema com t_1 livre).

Teorema 5.3. *Suponha que as condições (5.1)-(5.2) do Problema 5.1 com t_1 fixo sejam válidas. Seja $x^o(t)$, $u^o(t)$, t_1 uma solução do Problema 5.4. Então existem $\psi(t)$ e λ_0 simultaneamente não nulos, tais que:*

$$\frac{d\psi(t)}{dt} = -\varphi_x^*(x^o(t), u^o(t), t)\psi(t) + \lambda_0 \Phi_x(x^o(t), u^o(t), t) \ q.t.p. \ t_0 \leq t \leq t_1, \tag{5.37}$$

$$H(x^o(t), u^o(t), \psi(t), t) = \max_{u \in M} H(x^o(t), u(t), \psi(t), t) \ q.t.p. \ t_0 \leq t \leq t_1, \tag{5.38}$$

$$H(x^o(t), u^o(t), \psi(t), t) = \lambda + \int_t^{t_1} [-\langle \varphi_t(x^o, u^o, t), \psi(t) \rangle + \lambda_0 \Phi_t(x^o, u^o, t)]dt, \tag{5.39}$$

onde λ é uma constante.

Para terminar esta seção vamos aplicar o principio do máximo para o problema clássico do cálculo de variações.

5.4 Aplicação: Problema clássico de calculo de variações.

Nesta seção vamos a considerar um problema básico do cálculo de variações e obter uma condição de otimalidade (equações de Euler-Lagrange). Isto com a finalidade de mostrar que os problemas de controle ótimo englobam os problemas do calculo de variações.

$$\min F(x, u) = \int_0^T \Phi(x(t), u(t), t)dt$$

$s.a$

$$\frac{dx(t)}{dt} = u(t); x(0) = c; x(T) = d.$$

Como, neste caso a equação variacional é não degenerado, podemos tomar $\lambda_0 = 1$. Portanto,

$$H(x^\circ, u, \psi, t) = u\psi - \Phi(x^\circ, u, t).$$

Onde $\psi(t)$ satisfaz a equação diferencial

$$\frac{d\psi}{dt} = \Phi_x(x, u^\circ, t)$$

dai,

$$\psi(t) = \psi(0) + \int_0^t \Phi_x(x, u^\circ, \tau) d\tau = \Phi_u(x^\circ, u^\circ, t).$$

Portanto o principio do máximo produz a seguinte desigualdade para todo u e t.

$$u^\circ\psi - \Phi(x^\circ, u^\circ, t) \geq u\psi - \Phi(x^\circ, u, t),$$

$$H(x^\circ, u^\circ, \psi, t) \geq H(x^\circ, u, \psi, t).$$

Alternativamente definimos,

$$E(x^\circ, u^\circ, u, t) = \Phi(x^\circ, u, t) - \Phi(x^\circ, u^\circ, t) + (u^\circ - u)\Phi_u(x^\circ, u^\circ, t),$$

(Função de Weierstrass), temos que:

$$E(x^\circ, u^\circ, u, t) \geq 0 \ \forall u, t. \tag{5.40}$$

Este é conhecido como a condição necessária de extremo forte (condição de Weierstrass) do cálculo de variações, para mais detalhes veja [26]. O Princípio do Máximo é portanto uma generalização natural da condição de Weierstrass para o problema de Controle Ótimo.

CAPÍTULO 6

CONDIÇÕES SUFICIENTES PARA PROBLEMAS DE CONTROLE

Nos Capítulos [4] e [5] estudamos as condições necessárias de otimalidade para certos tipos de problemas de controle ótimo que são dadas pelo Princípio de Máximo de Pontryagin. Em relação as condições suficientes podemos citar trabalhos envolvendo hipóteses de convexidade, ver [9], condições de segundo ordem ver, [34] e [43], e o método da função de verificação via teoria de Hamilton-Jacobi-Bellman, veja [4], [8], [41] entre outros.

Neste capítulo estudamos as condições suficientes sob hipóteses de convexidade e invexidade generalizada. Na primeira seção determinaremos sob quais hipóteses as condições necessárias de otimalidade para problemas de controle com tempos finais fixos e tempos finais livres considerados nos capítulos [4] e [5] (Seção 5.2) respectivamente, são também suficientes. Na segunda seção vamos estudar a suficiência de um problema de controle com tempo final livre considerado na Seção 5.2 sob hipóteses de invexidade generalizada, que denominaremos PM-pseudo-invexidade.

6.1 Condições suficientes sob hipóteses de convexidade

Começamos considerando o problema de controle como no Capítulo 4:

$$(PCO)\begin{cases} \min F(x,u) = \displaystyle\int_0^T \Phi(x(t),u(t),t)dt \\ s.a \\ \qquad \dfrac{dx(t)}{dt} = \varphi(x(t),u(t),t) \text{ q.t.p. } 0 \leq t \leq T, \\ \qquad x(0) = c,\ x(T) = d, \\ \qquad u(t) \in M. \end{cases}$$

Com as mesmas hipóteses do Teorema 4.1.

Definição 6.1. *Dizemos que* $(x,u) \in C([0,T];\mathbb{R}^n) \times L_\infty([0,T];\mathbb{R}^r)$ *é um* ***processo factível*** *se* (x,u) *satisfaze as restrições do problema de controle (PCO).*

Definição 6.2. *Dizemos que* $(\bar{x},\bar{u}) \in C([0,T];\mathbb{R}^n) \times L_\infty([0,T];\mathbb{R}^r)$ *é um* ***processo extremal****, se existem* $\lambda_0 \geq 0$ *e* $\psi : [0,T] \to \mathbb{R}^n$ *satisfazendo as condições do Princípio de Máximo de Pontryagin para quase todo ponto* $t \in [0,T]$.

Definição 6.3. *Dizemos que* $(\bar{x},\bar{u}) \in C([0,T];\mathbb{R}^n) \times L_\infty([0,T];\mathbb{R}^r)$ *é um* ***processo ótimo*** *se* $(\bar{x},\bar{u})$ *é uma solução de (PCO), isto é,* $F(x,u) \geq F(\bar{x},\bar{u})$ *para todo processo factível* (x,u).

Teorema 6.1. *Seja* $(\bar{x},\bar{u})$ *um processo extremal normal de problema (PCO). Se* $H(x,u,\psi,t)$ *é côncava nas variáveis* x *e* u*, então* $(\bar{x},\bar{u})$ *é um processo ótimo.*

Demonstração. Por hipótese $H(x,u,\psi,t)$ é côncava em (x,u) para quase todo $t \in [0,T]$. Além disso, $\Phi(x(t),u(t),t)$ e $\varphi(x(t),u(t),t)$ são continuamente diferenciáveis em (x,u) (de classe C^1 nas variáveis x e u), portanto $H(x,u,t)$ é também de classe C^1, logo para cada par $(x,u),(\bar{x},\bar{u}) \in C([0,T];\mathbb{R}^n) \times L_\infty([0,T];\mathbb{R}^r)$ tem-se que:

$$\begin{aligned} H(x(t),u(t),t) - H(\bar{x}(t),\bar{u}(t),t) \quad &\leq \quad \langle H_x^*(\bar{x}(t),\bar{u}(t),t), x(t)-\bar{x}(t)\rangle + \\ &\quad\ \ \langle H_u^*(\bar{x}(t),\bar{u}(t),t), u(t)-\bar{u}(t)\rangle\,. \end{aligned} \tag{6.1}$$

Como $(\bar{x},\bar{u})$ é uma solução extremal de (PCO), então existem $\lambda_0 \geq 0$ $(\lambda_0 = 1)$ e $\psi : [0,T] \to \mathbb{R}^n$ satisfazendo (4.7) e (4.8) do Teorema 4.1, isto implica que,

$$H_u^*(\bar{x}(t),\bar{u}(t),t) = 0 \text{ q.t.p. } 0 \leq t \leq T.$$

Como a desigualdade é válida para quase todo $0 \leq t \leq T$ podemos integrar e a desigualdade (6.1) se preserva.

$$\int_0^T H(x(t),u(t),t)dt \leq \int_0^T \left[H(\bar{x}(t),\bar{u}(t),t) + \langle H_x^*(\bar{x}(t),\bar{u}(t),t), x(t)-\bar{x}(t)\rangle\right]dt. \quad (6.2)$$

De (6.2), da Observação 4.2, (4.7) e usando integração por partes temos

$$\begin{aligned}
&\int_0^T \left[\Phi(\bar{x}(t),\bar{u}(t),t) - \Phi(x(t),u(t),t)\right]dt \\
\leq\ & \int_0^T \langle \varphi(\bar{x}(t),\bar{u}(t),t) - \varphi(x(t),u(t),t), \psi(t)\rangle\, dt \\
&+ \int_0^T \langle \varphi_x^*(\bar{x}(t),\bar{u}(t),t)\psi(t) - \Phi_x(\bar{x}(t),\bar{u}(t),t), x(t)-\bar{x}(t)\rangle\, dt \\
=\ & \int_0^T \langle \bar{x}'(t) - x'(t), \psi(t)\rangle\, dt + \int_0^T \langle \psi'(t), \bar{x}(t)-x(t)\rangle\, dt \\
=\ & \int_0^T \langle \bar{x}'(t) - x'(t), \psi(t)\rangle\, dt + \langle \bar{x}(t)-x(t), \psi(t)\rangle\,|_0^T - \int_0^T \langle \bar{x}'(t) - x'(t), \psi(t)\rangle\, dt = 0.
\end{aligned}$$

Assim, $F(x,u) \geq F(\bar{x},\bar{u})$ $\forall (x,u) \in C([0,T],\mathbb{R}^n) \times L_\infty([0,T],\mathbb{R}^r)$. Portanto, $(\bar{x},\bar{u})$ é uma solução ótima. □

Agora consideremos um problema de controle como na Seção 6.1 do Capítulo [5], problema de controle ótimo não-autônomo com tempo final livre sob as hipóteses do Teorema 5.2:

$$(PCO^1)\begin{cases}
\min F(x,u) = \displaystyle\int_{t_0}^{t_1} \Phi(x(t),u(t),t)dt \\
s.a \\
\qquad \dfrac{dx(t)}{dt} = \varphi(x(t),u(t),t) \text{ q.t.p. } t_0 \leq t \leq t_1, \\
\qquad x(t_0) = c,\ x(t_1) = d, \\
\qquad u(t) \in M \text{ q.t.p. } t_0 \leq t \leq t_1.
\end{cases}$$

Definição 6.4. *Dizemos que $(x,u,t_1) \in C([t_0,t_1];\mathbb{R}^n) \times L_\infty([t_0,t_1];\mathbb{R}^r) \times [t_0,+\infty)$ é um* ***processo factível*** *se (x,u,t_1) satisfaze as restrições do problema de controle (PCO^1).*

Definição 6.5. *Dizemos que $(\bar{x},\bar{u},t_1) \in C([t_0,t_1];\mathbb{R}^n) \times L_\infty([t_0,t_1];\mathbb{R}^r)$ é um* ***processo extremal****, se existem $\lambda_0 \geq 0$ e $\psi : [t_0,t_1] \to \mathbb{R}^n$ satisfazendo as condições do Princípio de Máximo de Pontryagin para quase todo ponto $t \in [t_0,t_1]$.*

Definição 6.6. *Dizemos que $(\bar{x},\bar{u},t_1) \in C([t_0,t_1];\mathbb{R}^n) \times L_\infty([t_0,t_1];\mathbb{R}^r) \times [t_0,+\infty)$ é um* ***processo ótimo*** *se $(\bar{x},\bar{u},t_1)$ é uma solução de (PCO^1), isto é, $F(x,u) \geq F(\bar{x},\bar{u})$ para todo processo factível (x,u,t_1).*

Teorema 6.2. *Seja $(\bar{x}, \bar{u}, t_1)$ um processo extremal normal de (PCO^1). Suponha que para cada $t \in [t_0, t_1]$ fixo, a aplicação $(x, u) \longmapsto H(x, u, \psi(t), t) + I_M(\bar{u})$ é côncava. Então $(\bar{x}, \bar{u})$ é uma solução ótima.*

Demonstração. Como $(\bar{x}, \bar{u}, t_1)$ um processo extremal normal de (PCO^1), então existem $\lambda_0 = 1$ e $\psi : [t_0, t_1] \to \mathbb{R}^n$ tais que satisfazem (5.35)-(5.38) do Teorema 5.2. De (5.38) temos que

$$H(\bar{x}(t), \bar{u}(t), \psi(t), t) = \max_{u \in M} H(\bar{x}(t), u(t), \psi(t), t)$$
$$\Longleftrightarrow -H(\bar{x}(t), \bar{u}(t), \psi(t), t) = \min_{u \in M} -H(\bar{x}(t), u(t), \psi(t), t).$$

Como $-H$ atinge seu mínimo sob M em $u = \bar{u}(t)$ e $-H$ é convexo na variável u, segue do Teorema C.2, apêndice C que:

$$\begin{aligned} 0 &\in \partial_u \left[-H(\bar{x}(t), u(t), \psi(t), t)\right] + N_M(\bar{u}(t)) \\ &= \partial_u \left[-H(\bar{x}(t), u(t), \psi(t), t)\right] + \partial_u I_M(\bar{u}(t)) \\ &= \partial_u \left[-H(\bar{x}(t), \bar{u}(t), \psi(t), t) + I_M(\bar{u}(t))\right] \text{ q.t.p. } t_0 \leq t \leq t_1. \end{aligned} \tag{6.3}$$

Na segunda igualdade usamos a Proposição C.4 e na terceira igualdade o Teorema C.4, com I_M sendo função indicadora. Logo, de (5.37) e (6.3) concluímos que

$$(\psi'(t), 0) \in \partial_{xu} \left[-H(\bar{x}(t), \bar{u}, \psi, t) + I_M(\bar{u}(t))\right].$$

Deste modo, para quase todo $t_0 \leq t \leq t_1$, segue da definição C.3 que

$$-H(x, u, \psi(t), t) + H(\bar{x}(t), \bar{u}(t), \psi(t), t) \geq \langle (\psi'(t), 0), (x - \bar{x}(t), u - \bar{u}(t)) \rangle \ \forall x \in \mathbb{R}^n, \ \forall u \in M.$$

Seja (x, u, t_1') um processo factível para (PCO^1). Tem-se, que para quase todo $t_0 \leq t \leq t_1$, que

$$\begin{aligned} \Phi(x(t), u(t), t) &\geq \langle \varphi(x(t), u(t), t), \psi(t) \rangle - \langle \varphi(\bar{x}(t), \bar{u}(t), t), \psi(t) \rangle \\ &\quad + \Phi(\bar{x}(t), \bar{u}(t), t) + \langle \psi'(t), x(t) - \bar{x}(t) \rangle \\ &= \langle \varphi(x(t), u(t), t) - \varphi(\bar{x}(t), \bar{u}(t), t), \psi(t) \rangle + \Phi(\bar{x}(t), \bar{u}(t), t) \\ &\quad + \langle \psi'(t), x(t) - \bar{x}(t) \rangle. \end{aligned}$$

Integrando esta última desigualdade temos

$$\begin{aligned}
\int_{t_0}^{t_1} \Phi(x(t), u(t), t)dt \;\; &\geq \;\; \int_{t_0}^{t_1} \Phi(\bar{x}(t), \bar{u}(t), t)dt + \int_{t_0}^{t_1} \langle x'(t) - \bar{x}'(t), \psi(t) \rangle \, dt \\
&\quad + \int_{t_0}^{t_1} \langle \psi'(t), x(t) - \bar{x}(t) \rangle \, dt \\
&= \int_{t_0}^{t_1} \Phi(\bar{x}(t), \bar{u}(t), t)dt + \langle x(t) - \bar{x}(t), \psi(t) \rangle \,|_{t_0}^{t_1} \\
&\quad - \int_{t_0}^{t_1} \langle x(t) - \bar{x}(t), \psi'(t) \rangle \, dt + \int_{t_0}^{t_1} \langle \psi'(t), x(t) - \bar{x}(t) \rangle \, dt \\
&= \int_{t_0}^{t_1} \Phi(\bar{x}(t), \bar{u}(t), t)dt \\
&\Longrightarrow F(x, u) \geq F(\bar{x}, \bar{u}).
\end{aligned}$$

Portanto, $(\bar{x}, \bar{u}, t_1)$ é uma solução ótima. □

6.2 Condições suficientes sob hipóteses de invexidade generalizada

Nesta seção introduzimos a noção de PM-pseudo-invexidade, noção que foi extendida do contexto da programação matemática para problemas de controle ótimo escalar (PCO[1]), a qual é essencial para mostrar que as condições necessárias de otimalidade estabelecidas pelo Teorema 5.2 são também suficientes. É possível mostrar que estas condições necessárias de otimalidade são suficientes se, e somente se, o problema é PM-pseudo-invexo.

Definição 6.7. *(Invexidade). Seja $\phi : S \subset \mathbb{R}^k \to \mathbb{R}$ uma função diferenciável, onde S é um conjunto aberto não vazio. Dizemos que ϕ é **invexa** se existe uma função $\eta : S \times S \to \mathbb{R}^n$ tal que*

$$\phi(x) - \phi(\bar{x}) \geq \nabla\phi(\bar{x})^T \eta(x, \bar{x}) \;\; \forall x, \bar{x} \in S.$$

Claramente que se $\phi(\cdot)$ é convexo, $\phi(\cdot)$ é invexo com $\eta(x, \bar{x}) = x - \bar{x}$.

Definição 6.8. *(Pseudo-invexidade). Uma função $\phi : S \subset \mathbb{R}^k \to \mathbb{R}$ é dita **pseudo-invexa**, se existe $\eta : S \times S \to \mathbb{R}^n$ tal que para todo $x, \bar{x} \in S$,*

$$\eta(x, \bar{x})^T \nabla\phi(\bar{x}) \geq 0 \Rightarrow \phi(x) - \phi(\bar{x}) \geq 0.$$

Agora vamos definir o conceito PM-pseudo-invexidade para um problema de controle. Logo provaremos que sob essa hipótese, a condição necessária de otimalidade é também suficiente.

Definição 6.9. *Dizemos que o Problema (PCO^1) é **PM-pseudo-invexo** se para cada par* (x, u, t_1), $(\bar{x}, \bar{u}, t_1') \in C([t_0, t_1]; \mathbb{R}^n) \times L_\infty([t_0, t_1]; \mathbb{R}^r) \times [t_0, +\infty)$ *de processos factíveis tais que* $F(x, u) < F(\bar{u}, \bar{x})$, *existem funções* $\eta : [t_0, t_1] \longrightarrow \mathbb{R}^n$ *e* $\nu : [t_0, t_1] \longrightarrow \mathbb{R}$ *verificando*

$$\eta'(t) = \varphi_x^*(\bar{x}(t), \bar{u}(t), t)\eta(t) + \varphi_t(\bar{x}(t), \bar{u}(t), t)\nu(t) + \varphi(\bar{x}(t), \bar{u}(t), t)\nu'(t) + \Delta\varphi(\bar{x}, u, t) \quad (6.4)$$

em quase todo ponto $t_0 \leq t \leq t_1$,

$$\eta(t_0) = \eta(t_1) = 0; \;\; \nu(t_0) = 0, \;\; \nu(t_1) \in \mathbb{R}, \quad (6.5)$$

$$\int_{t_0}^{t_1} [\langle \Phi_x(\bar{x}(t), \bar{u}(t), t), \eta(t)\rangle + \Phi_t(\bar{x}(t), \bar{u}(t), t)\nu(t) + \Phi(\bar{x}(t), \bar{u}(t), t)\nu'(t) + \Delta\Phi(\bar{x}, u, t)]\, dt < 0, \quad (6.6)$$

onde,

$$\Delta\Phi(\bar{x}(t), u(t), t) := \Phi(\bar{x}(t), u(t), t) - \Phi(\bar{x}(t), \bar{u}(t), t) \;\; e$$

$$\Delta\varphi(\bar{x}(t), u(t), t) := \varphi(\bar{x}(t), u(t), t) - \varphi(\bar{x}(t), \bar{u}(t), t)$$

são definidos como em Vinter [41], isto com a finalidade de linearizar o sistema dinâmico.

Observação 6.1. *Na definição acima as funções* η *e* ν *dependem tanto de* (x, u, t_1') *como de* $(\bar{x}, \bar{u}, t_1)$.

Teorema 6.3. *Seja* $(\bar{x}, \bar{u}, t_1)$ *um processo extremal normal do Problema (PCO^1). Suponha o problema é MP-pseudo-invexo, então* $(\bar{x}, \bar{u}, t_1)$ *é um processo ótimo.*

Demonstração. Suponha por absurdo que $(\bar{x}, \bar{u}, t_1)$ não seja um processo ótimo, então existe um processo factível $(x, u, t_1') \in C([t_0, t_1]; \mathbb{R}^n) \times L_\infty([t_0, t_1]; \mathbb{R}^r) \times [t_0, +\infty)$ tais que $F(x, u) < F(\bar{x}, \bar{u})$. Como (PCO^1) é PM-pseudo-invexo, existem $\eta : [t_0, t_1] \to \mathbb{R}^n$ e $\nu : [t_0, t_1] \to \mathbb{R}$ verificando (6.4)-(6.6).
De (6.4) temos que

$$\int_{t_0}^{t_1} \langle \eta'(t), \psi(t)\rangle dt$$

$$= \int_{t_0}^{t_1} \langle \varphi_x(\bar{x}(t), \bar{u}(t), t)\eta(t) + \varphi_t(\bar{x}(t), \bar{u}(t), t)\nu(t) + \varphi(\bar{x}(t), \bar{u}(t), t)\nu'(t) + \Delta\varphi(\bar{x}, u, t), \psi(t)\rangle dt. \quad (6.7)$$

Agora, de (6.6) e (6.7) segue que

$$\begin{aligned}
\Sigma \;:=\; & \int_{t_0}^{t_1} \left[\langle \Phi_x(\bar{x}(t), \bar{u}(t), t), \eta(t)\rangle + \Phi_t(\bar{x}(t), \bar{u}(t), t)\nu(t) + \Phi(\bar{x}(t), \bar{u}(t), t)\nu'(t)\right] dt \\
& + \int_{t_0}^{t_1} \Delta\Phi(\bar{x}(t), u(t), t) dt - \int_{t_0}^{t_1} \langle \Delta\varphi(\bar{x}(t), u(t), t), \psi(t)\rangle dt + \int_{t_0}^{t_1} \langle \eta'(t), \psi(t)\rangle dt \\
& - \int_{t_0}^{t_1} \langle \varphi_x(\bar{x}(t), \bar{u}(t), t)\eta(t) + \varphi_t(\bar{x}(t), \bar{u}(t), t)\nu(t) + \varphi(\bar{x}(t), \bar{u}(t), t)\nu'(t), \psi(t)\rangle \, dt \\
& < 0. \qquad (6.8)
\end{aligned}$$

Por outro lado,

$$\begin{aligned}
\Sigma \;=\; & \int_{t_0}^{t_1} \langle \Phi_x(\bar{x}(t), \bar{u}(t), t) - \varphi_x(\bar{x}(t), \bar{u}(t), t)\psi(t), \eta(t)\rangle \, dt \\
& + \int_{t_0}^{t_1} \left[\Phi_t(\bar{x}(t), \bar{u}(t), t) - \langle \varphi_t(\bar{x}(t), \bar{u}(t), t), \psi(t)\rangle\right] \nu(t) dt \\
& + \int_{t_0}^{t_1} \left[\Phi(\bar{x}(t), \bar{u}(t), t) - \varphi(\bar{x}(t), \bar{u}(t), t)\psi(t)\right] \nu'(t) dt \\
& + \int_{t_0}^{t_1} \langle \eta'(t), \psi(t)\rangle dt - \int_{t_0}^{t_1} \left[\langle \varphi(\bar{x}(t), u(t), t), \psi(t)\rangle - \Phi(\bar{x}(t), u(t), t)\right] dt \\
& + \int_{t_0}^{t_1} \left[\langle \varphi(\bar{x}(t), \bar{u}(t), t), \psi(t)\rangle - \Phi(\bar{x}(t), \bar{u}(t), t)\right] dt \\
=\; & \int_{t_0}^{t_1} \langle \psi'(t), \eta(t)\rangle \, dt + \int_{t_0}^{t_1} r'(t)\nu(t) dt - \int_{t_0}^{t_1} H(\bar{x}(t), \bar{u}(t), \psi(t), t)\nu'(t) dt + \\
& \int_{t_0}^{t_1} \langle \eta'(t), \psi(t)\rangle \, dt + \int_{t_0}^{t_1} \left[H(\bar{x}(t), \bar{u}(t), \psi(t), t) - H(\bar{x}(t), u(t), \psi(t), t)\right] dt, \qquad (6.9)
\end{aligned}$$

onde $r(t)$ é solução de

$$\begin{aligned}
& r'(t) = \Phi_t(\bar{x}(t), \bar{u}(t), t) - \langle \varphi_t(\bar{x}(t), \bar{u}(t), t), \psi(t)\rangle \ \text{ q.t.p. } t_0 \le t \le t_1, \\
& r(t_1) = 0.
\end{aligned} \qquad (6.10)$$

Além disso, de (5.35) e (6.10) tem-se para quase todo ponto $t_0 \le t \le t_1$ que

$$\begin{aligned}
-\, r(t) \;=\; & \int_{t}^{t_1} \left[\Phi_t(\bar{x}(\tau), \bar{u}(\tau), \tau) - \langle \varphi_t(\bar{x}(\tau), \bar{u}(\tau), \tau), \psi(\tau)\rangle\right] dt \\
=\; & H(\bar{x}(t), \bar{u}(t), \psi(t), t). \qquad (6.11)
\end{aligned}$$

De (5.38) tem-se que o Hamiltoniano é maximizado em $\bar{u}(t)$, isto é,

$$H(\bar{x}(t), \bar{u}(t), \psi(t), t) - H(\bar{x}(t), u(t), \psi(t), t) \ge 0 \text{ q.t.p. } t_0 \le t \le t_1. \qquad (6.12)$$

Logo de (6.9), (6.12), usando integração por partes e tendo em conta (6.5) e (6.11) temos

$$\begin{aligned}
\Sigma &= \int_{t_0}^{t_1} \langle \psi'(t), \eta(t) \rangle \, dt + r(t)\nu(t)|_{t_0}^{t_1} - \int_{t_0}^{t_1} r(t)\nu'(t)dt \\
&\quad - \int_{t_0}^{t_1} H(\bar{x}(t), \bar{u}(t), \psi(t), t)\nu'(t)dt + \langle \eta(t), \psi(t) \rangle |_{t_0}^{t_1} - \int_{t_0}^{t_1} \langle \psi'(t), \eta(t) \rangle \, dt \\
&\quad + \int_{t_0}^{t_1} [H(\bar{x}(t), \bar{u}(t), \psi(t), t) - H(\bar{x}(t), u(t), \psi(t), t)] \, dt \\
&= - \int_{t_0}^{t_1} r(t)\nu'(t)dt - \int_{t_0}^{t_1} H(\bar{x}(t), \bar{u}(t), \psi(t), t)\nu'(t)dt \\
&\quad + \int_{t_0}^{t_1} [H(\bar{x}(t), \bar{u}(t), \psi(t), t) - H(\bar{x}(t), u(t), \psi(t), t)] \, dt \\
&= - \int_{t_0}^{t_1} r(t)\nu'(t)dt - \int_{t_0}^{t_1} -r(t)\nu'(t)dt \\
&\quad + \int_{t_0}^{t_1} [H(\bar{x}(t), \bar{u}(t), \psi(t), t) - H(\bar{x}(t), u(t), \psi(t), t)] \, dt \geq 0,
\end{aligned}$$

que contradiz com (6.8). Portanto $(\bar{x}, \bar{u}, t_1)$ é um processo ótimo. □

O resultados a seguir é uma caraterização de um processo extremal.

Proposição 6.1. *Seja $(\bar{x}, \bar{u}, t_1)$ um processo factível de (PCO^1). Suponha que dado $u(t) \in M$ q.t.p. $t_0 \leq t \leq t_1$, para todo par (η, ν) tal que*

$$\eta'(t) = \varphi_x^*(\bar{x}(t), \bar{u}(t), t)\eta(t) + \varphi_t(\bar{x}(t), \bar{u}(t), t)\nu(t) + \varphi(\bar{x}(t), \bar{u}(t), t)\nu'(t) + \Delta\varphi(\bar{x}(t), u(t), t) \tag{6.13}$$

em quase todo ponto $t_0 \leq t \leq t_1$,

$$\eta(t_0) = \eta(t_1) = 0, \ \nu(t_0) = 0, \ \nu(t_1) \in \mathbb{R}, \tag{6.14}$$

tem-se

$$\int_{t_0}^{t_1} [\langle \Phi_x(\bar{x}(t), \bar{u}(t), t), \eta(t) \rangle + \Phi_t(\bar{x}(t), \bar{u}(t), t)\nu(t) + \Phi(\bar{x}(t), \bar{u}(t), t)\nu'(t) + \Delta\Phi(\bar{x}, u, t)] \, dt \geq 0. \tag{6.15}$$

Então $(\bar{x}, \bar{u}, t_1)$ é um processo extremal.

Demonstração. Consideremos o problema de controle auxiliar (PCA):

$$(PCA)\begin{cases} \min \mathcal{F}(\eta,\nu,\xi,u) = \displaystyle\int_{t_0}^{t_1} \Lambda(\eta(t),\nu(t),\xi(t),u(t))dt \\ s.a \\ \qquad \eta'(t) = \Theta(\eta(t),\nu(t),\xi(t),u(t)) \text{ q.t.p. } t_0 \le t \le t_1, \\ \qquad \eta(t_0) = \eta(t_1) = 0, \\ \qquad \nu'(t) = \xi(t), \nu(t_0) = 0, \nu(t_1) \in \mathbb{R}, \\ \qquad \xi(t) \in \mathbb{R},\ u(t) \in M \text{ q.t.p. } t_0 \le t \le t_1, \end{cases}$$

onde

$$\Lambda(\eta,\nu,\xi,u) := \langle \Phi_x(\bar{x}(t),\bar{u}(t),t),\eta(t)\rangle + \Phi_t(\bar{x}(t),\bar{u}(t),t)\nu(t) + \Phi(\bar{x}(t),\bar{u}(t),t)\xi(t) + \Delta\Phi(\bar{x},u,t),$$

$$\Theta(\eta,\nu,\xi,u) := \varphi_x(\bar{x}(t),\bar{u}(t),t)\eta(t) + \varphi_t(\bar{x}(t),\bar{u}(t),t)\nu(t) + \varphi(\bar{x}(t),\bar{u}(t),t)\xi(t) + \Delta\varphi(\bar{x},u,t).$$

Observe que $(\bar{\eta},\bar{\nu},\bar{\xi},\bar{u}) = (0,0,0,\bar{u})$ é um processo factível. Além disso, $\mathcal{F}(\bar{\eta},\bar{\nu},\bar{\xi},\bar{u}) = 0 \le \mathcal{F}(\eta,\nu,\xi,\bar{u})$ para todo processo factível, por hipótese. Logo $(\bar{\eta},\bar{\nu},\bar{\xi},\bar{u})$ é um processo ótimo. Pelo princípio de Máximo, Teorema 5.2, existem $\lambda_0 \ge 0$ e $\psi : [t_0,t_1] \to \mathbb{R}^{n+1}$, $\psi := (\psi_1,\psi_2)$ tais que

(i) $(\lambda_0,\psi) \ne (0,0)$,

(ii) a) $\psi_1'(t) = -\varphi_x^*(\bar{x}(t),\bar{u}(t),t)\psi_1(t) + \lambda_0\Phi_x(\bar{x}(t),\bar{u}(t),t)$ q.t.p. $t_0 \le t \le t_1$,

b) $\psi_2'(t) = -\varphi_t^*(\bar{x}(t),\bar{u}(t),t)\psi_1(t) + \lambda_0\Phi_t(\bar{x}(t),\bar{u}(t),t)$ q.t.p. $t_0 \le t \le t_1$,

(iii)

$$\max_{u,\xi\in M\times\mathbb{R}} \{ \langle \psi_1(t), \varphi_x(\bar{x}(t),\bar{u}(t),t)\bar{\eta}(t) + \varphi_t(\bar{x}(t),\bar{u}(t),t)\bar{\nu}(t) + \varphi(\bar{x}(t),\bar{u}(t),t)\xi + \Delta\varphi(\bar{x}(t),u,t)\rangle + \psi_2(t)\xi(t) - \lambda_0\Phi(\bar{x}(t),\bar{u}(t),t) - \lambda_0\Delta\Phi(\bar{x}(t),u(t),t) \} = 0$$

$\Leftrightarrow$

$$\max_{u,\xi\in M\times\mathbb{R}} \{\langle \psi_1(t), \varphi(\bar{x}(t),\bar{u}(t),t)\xi + \Delta\varphi(\bar{x}(t),u,t)\rangle + \psi_2(t)\xi - \lambda_0\Phi(\bar{x}(t),\bar{u}(t),t) - \lambda_0\Delta\Phi(\bar{x}(t),u,t) \} = 0$$

$\Leftrightarrow$

$$\langle \psi_1(t), \varphi(\bar{x}(t),\bar{u}(t),t)\xi + \Delta\varphi(\bar{x}(t),u,t)\rangle + \psi_2(t)\xi - \lambda_0\Phi(\bar{x}(t),\bar{u}(t),t)\xi - \lambda_0\Delta\Phi(\bar{x}(t),u,t) \le 0,$$

para todo $u \in M$, $\forall \xi$.

(iv) $\psi_2(t_1) = 0$.

Em particular, para $\xi = 1$ e $u = \bar{u}(t)$ temos

$$\langle \psi_1(t), \varphi(\bar{x}(t), \bar{u}(t), t)\rangle + \psi_2(t) - \lambda_0 \Phi(\bar{x}(t), \bar{u}(t), t) \leq 0 \text{ q.t.p. } t_0 \leq t \leq t_1,$$

para $\xi = -1$ e $u = \bar{u}(t)$, obtemos a desigualdade oposta. Logo

$$H(\bar{x}(t), \bar{u}(t), \psi, t) = -\psi_2(t) \text{ q.t.p. } t_0 \leq t \leq t_1.$$

Desta última igualdade, integrando em b) e de iv) obtemos (5.35).
Para $\xi = 0$,

$$\langle \psi_1(t), \Delta\varphi(\bar{x}(t), u, t)\rangle - \lambda_0 \Delta\Phi(\bar{x}(t), u, t) \leq 0 \ \forall u \in M$$

$$\Leftrightarrow \langle \psi_1(t), \varphi(\bar{x}(t), u, t)\rangle - \lambda_0 \Phi(\bar{x}(t), u, t) \leq \langle \psi_1(t), \varphi(\bar{x}(t), \bar{u}(t), t)\rangle - \lambda_0 \Phi(\bar{x}(t), \bar{u}(t), t)$$

$$\Leftrightarrow H(\bar{x}(t), u, \psi_1(t), t) \leq H(\bar{x}(t), \bar{u}(t), \psi_1(t), t) \ \forall u \in M.$$

Portanto,

$$\max_{u \in M} H(\bar{x}(t), u, \psi_1(t), t) = H(\bar{x}(t), \bar{u}(t), \psi_1(t), t) \text{ q.t.p. } t_0 \leq t \leq t_1.$$

Donde (5.38) é satisfeita. De (iii)-(a), segue que (5.37) vale. Claro que $(\lambda_0, \psi_1) \neq (0, 0)$, pois caso contrário teríamos $\psi_2'(t) = 0$, i.e., $\psi_2(t) = cst := 0$ q.t.p. $t \in [t_0, t_1]$. Logo $(\lambda_0, \psi_1, \psi_2) = (0, 0, 0)$, contradizendo (i).
Logo são satisfeitas (5.37)-(5.35) do Teorema 5.2. Portanto, $(\bar{x}, \bar{u}, t_1)$ é um processo extremal. □

Teorema 6.4. *Se (PCO^1) é tal que todo processo extremal é um processo ótimo, então (PCO^1) é PM-pseudo-invexo.*

Demonstração. Suponha o contrário, que (PCO^1) não seja PM-pseudo-invexo. Então existem um par de processos factíveis $(x, u, t_1'), (\bar{x}, \bar{u}, t_1) \in C([0, T], \mathbb{R}^n) \times L_\infty([0, T], \mathbb{R}^r) \times [t_0, +\infty)$ tais que $F(x, u) < F(\bar{x}, \bar{u})$ e para qualquer η e ν verificando (6.4)-(6.5), a desigualdade (6.6) não se verifica. Pela Proposição 6.1, $(\bar{x}, \bar{u}, t_1)$ é um processo extremal. Mas $F(x, u) < F(\bar{x}, \bar{u})$ implica que $(\bar{x}, \bar{u}, t_1)$ não é processo ótimo, isto contradiz a hipótese que todo processo extremal é processo ótimo. Portanto (PCO^1) é PM-pseudo-invexo. □

O seguinte teorema segue diretamente dos Teoremas 6.3 e 6.4, e mostra que os problemas de controle PM-pseudo-invexos são a classe mais ampla com a propriedade que todo processo extremal é ótimo.

Teorema 6.5. *Suponha que todo processo extremal de (PCO^1) é normal. (PCO^1) é PM-pseudo-invexo se, somente se, todo processo extremal é ótimo.*

CONCLUSÕES

O objetivo principal desta dissertação foi discutir condições necessárias e suficientes de otimalidade para problemas de controle ótimo.

No Capítulo 4 estudamos um problema de controle com restrição sobre o controle, envolvendo hipóteses de diferenciabilidade a respeito das variáveis de estado e controle dos funcionais de custo e de estado. Logo fizemos uso do formalismo de Dubovitskii-Milyutin para a derivação de condições necessárias de otimalidade.

No Capítulo 5 estudamos um problema de controle autônomo, Problema 5.1, sob hipóteses mais fracas no controle, com tempo final livre e tirando a diferenciabilidade no controle dos funcionais, onde condição necessária é dado pelo Teorema 4.1, em cuja prova se reduz ao caso tratado no Capítulo 4, introduzindo um problema auxiliar. Como variante a este problema, estudamos problemas de controle ótimo não-autônomos.

Finalmente no Capítulo 6. Na Seção 6.1 obtivemos uma caraterização para a condição suficiente de otimalidade para os Problemas 4.1 e 5.4, impondo uma hipótese de concavidade no Hamiltoniano com respeito às variáveis de estado e controle. E, na Seção 6.2 estudamos a derivação condição suficientes de otimalidade para problemas de controle não-autônomos com tempo final livre, introduzindo a noção de PM-pseudo-invexidade. Mostramos que a hipótese de PM-pseudo-invexidade é uma condição necessária e suficiente de otimalidade para (PCO^1), isto é, se todo processo factível é um processo ótimo se, e somente se, o problema é PM-pseudo-invexo.

Como trabalho futuro, um grande desafio é estudar as condições de otimalidade para problemas de controle não suave via formalismo Dubovitski-Milyutin. Assim como também o estudo problemas de controle multiobjetivo usando esse mesmo formalismo.

REFERÊNCIAS BIBLIOGRÁFICAS

[1] Alekseev, V. M., Tikhomirov, V. M., Fomin, S. V. *Optimal Control.* Springer-Moscow, Moscow, 1987.

[2] Bartle, R. *The Element of Integration and Lebesgue Measure.* Wiley, 1995.

[3] Bazaara, M. S., Sherali, C. M., Shetty, C. M. *Nonlinear Programming. Theory and Algorihms.* Jhon Wiley and Sons, Inc. New York, 1993.

[4] Bertsekas, D. *Dynamic Programming and Optimal Control.* Anthena Scientific, Massachusetts, 2005.

[5] Boyd, S., Vandenberghe, L. *Convex Optimization.* Cambridge University Press, Los Angeles, 2009.

[6] Brandão, M. A., Rojas Medar, M. A. Nonsmooth continuous-time optimization problems: necessary conditions. *Comp. Math. with Appl. 41* (2001), p. 1477–1486.

[7] Brezis, H. *Functional Analysis, Sobolev Spaces and Partial Differential Equations.* Springer New York Dordrecht Heidelberg London, New York, 2011.

[8] Capuzo, I., Bardi, M. *Optimal Control and Viscosity Solutions of Hamilton-Jacobi-Bellman Equations.* Birkhäuser, Boston, 2008.

[9] Cerda, E. *Optimización Dinámica.* Editora Prentice, Madrid, 2006.

[10] Challco, C. Y. Controle Ótimo via inclusões diferenciais. Master's thesis, Universidade Estadual de Campinas, 2000.

[11] Christensen, G. S., El-Haway, M. E., Soliman, S. A. *Optimal Control Aplications in Electric Power Systems.* Plenum, New Jersey, 1987.

[12] Clarke, F. *Functional Analysis, Calculus of Variations and Optimal Control. Graduate Texts in Mathematics.* Springer, New York, 2013.

[13] Craven, B. D. Invex functions and constrained local minima. *Bull. Austral. Math. Soc. 24* (1981), p. 357–366.

[14] Craven, B. D. *Control And Optimization.* Chapman & Hall, London, 1995.

[15] de Oliveira, V. A., Silva, G. N. New optimality conditions for nonsmooth control problems. *J. Glob. Optim. 57* (2013), p. 1465–1484.

[16] de Oliveira, V. A., Silva, G. N. On sufficient optimality conditions for multiobjective control problems. *J. Global Optim. (online), DOI 10.1007/s10898-015-0351-y* (2015).

[17] de Oliveira, V. A., Silva, G. N., Rojas-Medar, M. A. A class of multiobjective control problems. *Optim. Control Appl. Meth. 30* (2009), p. 77–86.

[18] de Oliveira, V. A., Silva, G. N., Rojas-Medar, M. A. Kt-invexity in optimal control problems. *Nonlinear Anal.-Theory Methods Appl. 71* (2009), p. 4790–4797.

[19] Fernandez, P. J. *Medida e Integração.* Coleção Projeto Euclides, 2007.

[20] Flett, T. M. *Diferential Analisys.* Cambridge University Press, London, 1908.

[21] Folland, G. *Real Analysis:Modern Techniques and their Applications.* Willey, 1999.

[22] Girsanov, I. V. *Lecture on Mathematical Theory of Extremum Problems.* Springer-Verlag, Berlin, Heidelberg, New York, 1972.

[23] Hanson, M. A. On sufficiency of the kuhn-tucker conditions. *Journal of Mathematical Analysis and Aplications 80* (1981), p. 545–550.

[24] Ioffe, A. D. Nonsmooth analysis:differential calculus of nondifferentiable mapping. *Trans. Amer. Soc. 266* (1981), p. 1–56.

[25] Ioffe, A. D. The maximum principle for abnormal optimal control problems. *Soviet Math. Dokl. 37* (1988), p. 134–231.

[26] Ioffe, A. D., M., T. V. *Theory of Extremal Problems.* North-Holland Publishing Company, New York, 1979.

[27] Kantorovich, L. V., Akilov, G. P. *Functional Analysis.* Nauka, 1982.

[28] Kreyszig, E. *Introductory Functional Analysis whit Aplicactions.* John Wiley and Sons, New York, 1978.

[29] Ledzewicz-Kowalewska, U. Aplication of some specification of the bubovitskii-milyutin method to the optimal control problem. *Non-linear Analisys. Methods and Application 12 N. 2* (1988), p. 101–108.

[30] Ledzewicz-Kowalewska, U. Extension of the local maximum principle to abnormal optimal control problem. *Non-linear Analisys. Theory and Application 77 No. 3* (1993), p. 661–681.

[31] Ledzewicz-Kowalewska, U. On abnormal control problem with mixed equality and inequality constraint. *Journal of Mathematical Analisys and Aplications 173* (1993), p. 18–42.

[32] Macki, J., Strauss, A. *Introduction to Optimal Control Theory.* Springer-Verlag, New York, 1982.

[33] Martin, D. H. The essence of invexity. *Journal of Optimization Theory and Aplications 47* (1985), p. 65–76.

[34] Maurer, H., Pickenhan, S. Second-order sufficient conditions optimal control problems with mixed control-state constraints. *Journal of Optimization Theory and Aplications 86, n. 3* (1995), p. 649–667.

[35] Mond, B., Husain, I. Sufficient optimality criteria and duality for variational problems with generalized invexity. *J. Austral Mat. Soc. Ser. B31* (1989), p. 108–121.

[36] Pinch, E. R. *Optimal Control and the Calculus of Variations.* United States by Oxford University Press Inc., New York, 1993.

[37] Pontryagin, L. S., Bontyanskki, R. G., Gamkrelidze, R. V., Mishchenko, E. F. *The Mathematical Theory of Optimal Processes.* Jonh Wiley & Sons, New York, 1962.

[38] Rojas-Medar, M. A., Brandão, M. A., Silva, G. N. Nonsmooth continuous-time optimization problems: necessary conditions. *J. Math. Anal. Appl. 227* (1998), p. 305–318.

[39] Shone, R. *Economic Dynamics.* Cambridge University Press, New York, 2002.

[40] Swam, G. W. *Applications of Optimal Control Theory in Biomedicine.* Marcel Dekker, New York, 1984.

[41] Vinter, R. *Optimal Control.* Birkhäuser, Boston, 2000.

[42] Vivanco, O. V. *Problema de Extremos Regulares y no Regulares, via formalismo de Dubovitskii-Milyutin. Aplicación a Problemas de Control Óptimo.* PhD thesis, Universidad de Sevilla, 2013.

[43] Zsolt, P., Vera, Z. Optimal control problems with set-valued control and state constraints. *SIAM Journal Optimization 346, n. 2* (2003), p. 334–358.

APÊNDICE A

CONJUNTO DE CANTOR DE MEDIDA POSITIVA

Esta seção trata sobre a construção de conjuntos de Cantor generalizados, no sentido de que é possível construir conjuntos que sejam perfeitos e nunca densos igual ao conjunto clássico de Cantor (de medida zero), mas neste caso com medida positiva. A diferença vem da construção, em lugar de omitir o terço médio de $[a,b]$ com $a=0$ e $b=1$, se remove qualquer intervalo centrado em $[a,b]$ de comprimento menor que $|b-a|$.

A construção de um conjunto tipo Cantor de medida positiva poder ser feita para qualquer $a \neq 0$ e $b \neq 1$. Por simplicidade tomaremos $a=0$ e $b=1$. Sejam $C=[0,1]$ e $0<\alpha<1$.

Passo 1. Retiramos de $[0,1]$ o intervalo $K_{1,1}=\left(\frac{1}{2}-\frac{1}{4}\alpha, \frac{1}{2}+\frac{1}{4}\alpha\right)$, cuja medida (medida de Lebesgue) é $\lambda(K_{1,1})=\frac{\alpha}{2}$, então o resto é

$$C_{1,1}=\left[0, \frac{1}{2}-\frac{1}{4}\alpha\right] \text{ e } C_{1,2}=\left[\frac{1}{2}+\frac{1}{4}\alpha, 1\right],$$

com

$$\lambda\left(C_{1,1}\right)=\lambda\left(C_{1,2}\right)=\frac{1}{2}\left(1-\frac{1}{2}\alpha\right).$$

Passo 2. Dos intervalos restantes excluímos os intervalos médios

$$K_{2,1}=\left(\frac{1}{4}-\frac{3}{16}\alpha, \frac{1}{4}-\frac{1}{16}\alpha\right),\ K_{2,2}=\left(\frac{3}{4}+\frac{1}{16}\alpha, \frac{3}{4}+\frac{3}{16}\alpha\right),$$

onde $\lambda\left(K_{2,1}\right)=\lambda\left(K_{2,2}\right)=\frac{\alpha}{8}$, então restam quatro intervalos

$$C_{2,1}=\left[0, \frac{1}{4}-\frac{3}{16}\alpha\right],\ C_{2,2}=\left[\frac{1}{4}-\frac{1}{16}\alpha, \frac{1}{2}-\frac{1}{4}\alpha\right],$$

$$C_{2,3} = \left[\frac{1}{2} + \frac{1}{4}\alpha, \frac{3}{4} + \frac{1}{16}\alpha\right] \text{ e } C_{2,4} = \left[\frac{3}{4} + \frac{3}{16}\alpha, 1\right]$$

onde

$$\lambda\left(C_{2,1}\right) = \lambda\left(C_{2,2}\right) = \lambda\left(C_{2,3}\right) = \lambda\left(C_{2,4}\right) = \frac{1}{4}\left(1 - \frac{\alpha}{2} - \frac{\alpha}{4}\right).$$

Passo 3. Dos quatro intervalos restantes do Passo 3. retiramos os intervalos médios de comprimentos iguais a $\dfrac{\alpha}{32}$, isto é, retiramos $K_{2,1}, K_{2,2}, K_{2,3}, K_{2,4}$ tais que

$$\lambda\left(K_{2,1}\right) = \lambda\left(K_{2,2}\right) = \lambda\left(K_{2,3}\right) = \lambda\left(K_{2,4}\right) = \frac{\alpha}{32}.$$

No seguinte passo, dos 8 intervalos restantes $C_{3,1}, C_{3,2}, \ldots, C_{3,8}$ cada um de comprimento $\dfrac{1}{8}\left(1 - \frac{\alpha}{12} - \frac{\alpha}{4} - \frac{\alpha}{8}\right)$, ou seja,

$$\lambda\left(C_{3,1}\right) = \lambda\left(C_{3,2}\right) = \cdots = \lambda\left(C_{3,8}\right) = \frac{1}{2^3}\left(1 - \frac{\alpha}{12} - \frac{\alpha}{4} - \frac{\alpha}{8}\right),$$

removemos os intervalos médios de comprimento $\frac{\alpha}{128}$ e construindo indutivamente temos:

Passo n. Depois de n etapas restam 2^n intervalos fechados $C_{n,1}, C_{n,2}, \ldots, C_{n,2^n}$ cada um de comprimento menor que $\frac{1}{2^n}$.

Sejam

$$A_n = \bigcup_{k=1}^{2^{n-1}} K_{n,k} \text{ e } B_n = \bigcup_{k=1}^{2^n} C_{n,k}$$

e definimos o conjunto

$$\mathcal{C} = \bigcap_{k=1}^{\infty} B_n.$$

A este conjunto assim definido chamaremos de conjunto de Cantor Generalizado.

Note que $(B_n)_n^\infty$ é uma sequência decrescente e $\lambda(B_1) < \infty$, então

$$\begin{aligned}
\lambda(\mathcal{C}) = \lim_{n\to\infty} \lambda(B_n) &= \lim_{n\to\infty} \left(1 - \lambda(A_n)\right) \\
&= 1 - \lim_{n\to\infty} \lambda\left(\bigcup_{k=1}^{2^{n-1}} A_{n,k}\right) \\
&= 1 - \lim_{n\to\infty} \sum_{k=1}^{2^{n-1}} \lambda(A_{n,k}) \\
&= 1 - \lim_{n\to\infty} \underbrace{\sum_{k=1}^{2^{n-1}} \frac{\alpha}{2^k}}_{=\alpha} \\
&= 1 - \alpha > 0.
\end{aligned}$$

Portando, obtemos um conjunto cuja medida é positiva e verifica-se as propriedades topológicas do conjunto de Cantor como, densidade, compacidade, não enumerabilidade, etc., veja [21].

APÊNDICE B

CONDIÇÃO DE NÃO DEGENERAÇÃO

Consideremos o problema de valor inicial

$$\frac{d\bar{x}(t)}{dt} = A(t)\bar{x}(t) + B(t)\bar{u}(t), \ \ \bar{x}(0) = 0, \tag{B.1}$$

onde $A(t)$ e $B(t)$ são matrizes $n \times n$ e $n \times r$, respectivamente, e $\bar{u}(t)$ é uma função arbitrária em $L_{\infty}^{(r)}[0, T]$.

Definição B.1. *(Condição de não degeneração $\mathbb{A}$). Dizemos que o sistema dado por (B.1) satisfaz a* ***condição de não degeneração*** *$\mathbb{A}$ ou simplesmente é* ***não degenerado****, se qualquer solução não nula $\psi(\cdot)$ da equação*

$$\frac{d\psi(t)}{dt} = -A^{*}(t)\psi(t) \tag{B.2}$$

satisfaze a condição $B^{*}(t)\psi(t) \neq 0$. Mais precisamente, $B^{*}(t)\psi(t)$ é não nulo sob um conjunto de medida positiva.

Observação B.1. *O sistema de equações diferenciais (B.1) que satisfaz a condição de não degeneração ($\mathbb{A}$) é dito completamente controlável.*

Existe outra condição de não degeneração, condição de não degeneração ($\mathbb{B}$).

Definição B.2. *(Condição de não degeneração $\mathbb{B}$). Seja $D \subset \mathbb{R}^n$ o conjunto de todos os vetores $\bar{x}(T)$, onde $\bar{x}(t)$ satisfaz a equação diferencial (B.1). Dizemos que o sistema (B.1) satisfaz a* ***condição de não degeneração ($\mathbb{B}$)*** *ou simplesmente é* ***não degenerado*** *se $D \equiv \mathbb{R}^n$.*

Lema B.1. *A condição de não degeneração ($\mathbb{A}$) implica a condição de não degeneração ($\mathbb{B}$).*

Demonstração. Suponha que a condição de não degeneração ($\mathbb{B}$) não seja válida, isto é, $D \neq R^n$. Como uma consequência da linearidade da equação, D é um subespaço donde segue que existe $a \in \mathbb{R}^n$, $a \neq 0$ ortogonal a D, ou seja, $\langle a, \bar{x}(T)\rangle = 0$ para todo $\bar{x}(t)$ que é solução da equação (B.1).
Considere agora o sistema

$$\begin{cases} \dfrac{d}{dt}\psi(t) = -A^*(t)\psi(t) \\ \psi(T) = a. \end{cases}$$

Como $a \neq 0$, segue que $\psi(t) \neq 0$. Então para qualquer solução $\bar{x}(t)$ de (B.1),

$$\begin{aligned} 0 &= \int_0^T \left\langle \frac{d}{dt}\psi(t) + A^*\psi(t), \bar{x}(t) \right\rangle dt \\ &= \langle \psi(T), \bar{x}(T)\rangle - \langle \psi(0), \bar{x}(0)\rangle - \int_0^T \left\langle \psi(t), \frac{d\bar{x}(t)}{dt} \right\rangle dt + \int_0^T \langle \psi(t), A(t)\bar{x}(t)\rangle\, dt \\ &= \langle a, \bar{x}(0)\rangle - \int_0^T \langle \psi(t), B(t)\bar{u}(t)\rangle\, dt \\ &= -\int_0^T \langle B^*(t)\psi(t), \bar{u}(t)\rangle\, dt. \end{aligned}$$

Logo para todo $\bar{u}(t) \in L_\infty([0,T];\mathbb{R}^r)$ temos

$$\int_0^T \langle B^*(t)\psi(t), \bar{u}(t)\rangle\, dt = 0,$$

mas isto é possível somente se $B^*(t)\psi(t) = 0$ q.t.p. $0 \leq t \leq T$ o que contradiz a hipótese. Portanto fica provado o lema. □

Agora, suponhamos que a condição de não degeneração $\mathbb{B}$ é satisfeita para o sistema (B.1), então $\Lambda'(x^o, u^o)$ é sobrejetora. Consideremos $(\bar{a}(t), b) \in E_1$ e tomando $z(t) \in C([0,T];\mathbb{R}^n)$ como no (3.4), tal que

$$z(t) = \bar{a}(t) + \int_0^t \varphi_x(x^o(\tau), u^o(\tau), \tau) z(\tau) d\tau \ \forall t \in [0,T], \tag{B.3}$$

como (B.3) é uma equação integral de de Volterra de segunda espécie, este possui uma única solução $z \in C([0,T];\mathbb{R}^n)$, para qualquer $\bar{a}(t) \in C([0,T];\mathbb{R}^n)$, veja [27] pág. 396. Note que a equação

$$\widehat{x}(t) - \int_0^t [\varphi_x(x^o(\tau), u^o(\tau), \tau)\widehat{x}(\tau) + \varphi_u(x^o(\tau), u^o(\tau), \tau)\widehat{u}(\tau)]\, d\tau = \bar{a}(t) \ \forall t \in [0,T],$$

tem em particular como solução $\widehat{u}(t) = 0$ e $\widehat{x}(t) = z(t)$, para qualquer $\bar{a} \in C([0,T];\mathbb{R}^n)$. Por outro lado, como o sistema (B.1) é não degenerado, podemos achar $\bar{u} \in L_\infty([0,T];\mathbb{R}^r)$ e $y \in C([0,T];\mathbb{R}^n)$ de modo que

$$\begin{aligned} &y'(t) = \varphi_x(x^o(t), u^o(t), t)y(t) + \varphi_u(x^o(t), u^o(t), t)\bar{u}(t),\\ &y(0) = 0,\\ &y(T) = b - z(T). \end{aligned} \tag{B.4}$$

Finalmente, colocando $\bar{x}(t) = y(t) + z(t)$ e $s = (x(\tau), u^o(\tau), \tau)$ por simplicidade, temos

$$\begin{aligned} &\Lambda'(x^o, u^o)(\bar{x}, \bar{u}) = \Big(\bar{x}(t) - \textstyle\int_0^t [\varphi_x(x^o(\tau), u^o(\tau), \tau)\bar{x}(\tau) + \varphi_u(x^o(\tau), u^o(\tau), \tau)\bar{u}(\tau)]\, d\tau, \bar{x}(T)\Big)\\ &= \Big(y(t) + z(t) - \textstyle\int_0^t [\varphi_x(s)(y(\tau) + z(\tau)) + \varphi_u(x^o(\tau), u^o(\tau), \tau)\bar{u}(\tau)]\, d\tau, \bar{x}(T)\Big). \end{aligned} \tag{B.5}$$

Assim por (B.5) temos que

$$\Lambda'(x^o, u^o)(\bar{x}, \bar{u}) = \left(y(t) + z(t) - \int_0^t \varphi_x(x^o(t), u^o(t), t)z(\tau)d\tau - y(\tau), b\right).$$

E por (B.4) obtemos $\Lambda'(x^o, u^o)(\bar{x}, \bar{u}) = (\bar{a}(t), b)$, ou seja, $\Lambda'(x^o, u^o)(\bar{x}, \bar{u})E = E_1$. Podemos dizer então, que a sobrejetividade do operador $\Lambda'(x^o, u^o)$ pode ser verificada através da condição de não degeneração $\mathbb{A}$, de onde basta encontrar a matriz fundamental A (matriz transição de estado) da equação

$$\frac{d\psi(t)}{dt} = -A^*\psi(t),$$

ou seja, basta resolver a equação matricial

$$\begin{aligned} &\tfrac{d\Theta(t)}{dt} = -A^*\Theta(t),\\ &\Theta(0) = I, \end{aligned}$$

onde I é a matriz identidade e Θ uma matriz $n \times n$, e depois achar a matriz $B^*\Theta(t)$. Se as colunas $\xi_i(t)$ da matriz $\Theta(t)$ são linearmente independentes como funções de t, ou seja, $\sum_{i=1}^n \lambda_i\xi_i(t) \neq 0$ para $\lambda = (\lambda, \ldots, \lambda_n) \neq 0$, então o sistema (B.1) é não degenerado. De fato, como qualquer solução $\psi(t)$ de (B.1) é da forma $\Theta(t)\lambda$, para $\lambda \neq 0$. Portanto

$$B^*(t)\Theta(t) = B^*(t)\Theta(t)\lambda = \sum_{i=1}^n \lambda_i\xi_i(t) \neq 0.$$

E como consequência, fica provado a afirmação da página 54.

APÊNDICE C

CÁLCULO SUBDIFERENCIAL

Notação: E espaço normado e $\mathbb{R}_\infty := \mathbb{R} \cup \{\infty\}$.

As definições e demostrações dos resultados apresentados em esta seção podem ser encontrados em Clarke [12] ou Vinter [41].

Definição C.1. *Seja S um subconjunto de E. O* ***cone normal*** *a S, $N_S(x)$, no ponto x é um subconjunto do espaço dual E^* definido por:*

$$N_S(x) := \{\xi \in E^* : \langle \xi, v \rangle \leq 0 \ \forall v \in T(S; x)\}.$$

Proposição C.1. *Seja $f : E \longrightarrow \mathbb{R}$ uma função diferenciável, e f atinge um mínimo sobre $A \subset E$ em um ponto x. Então temos*

$$-f'(x) \in N_A(x) \ \ e \ \ \langle f'(x), v \rangle \geq 0 \ \forall v \in T(A, x).$$

Definição C.2. *Seja $S \subset E$. Uma função $I_S : E \longrightarrow \mathbb{R}_\infty$ definida por*

$$I_S(x) = \begin{cases} 0, & se \ x \in S \\ \infty, & se \ x \notin S, \end{cases}$$

chama-se ***função indicadora***.

Definição C.3. *Seja $f : E \longrightarrow \mathbb{R}_\infty$ e seja $x \in dom f$. Um elemento $\xi \in E^*$ chama-se* ***subgradiente*** *de f em x, se satisfaze a seguinte desigualdade subdiferencial:*

$$f(y) - f(x) \geq \langle \xi, y - x \rangle, \ \ y \in E.$$

O conjunto de todos subgradientes de f em x é denotado por $\partial f(x)$.

Proposição C.2. *Seja $f : E \longrightarrow \mathbb{R}_\infty$ uma função convexa, e $x \in domf$. Então*

$$\partial f(x) = \{\xi \in E^* : f'(x,v) \geq \langle \xi, v \rangle \;\; \forall v \in E\}.$$

Proposição C.3. *Seja $f : E \longrightarrow \mathbb{R}_\infty$ uma função convexa com $x \in domf$. Se f é Gâteaux diferenciável, então $\partial f(x) = \{f'(x)\}$.*

Proposição C.4. *Seja $x \in S$, com S subconjunto convexo de E. Então $\partial I_S(x) = N_S(x)$.*

Teorema C.1. *Sejam $f, g : E \longrightarrow \mathbb{R}_\infty$ funções convexas tais que admitem um ponto em $domf \cap domg$ onde f é contínua. Então*

$$\partial\left(f+g\right)(x) = \partial f(x) + \partial g(x) \;\forall x \in domf \cap domg.$$

Teorema C.2. *Seja $f : E \longrightarrow \mathbb{R}_\infty$ uma função convexa e contínua, e seja A um subconjunto convexo de E. Então são equivalentes:*

(i) x é minimizador de f sob A.

(ii) $-\partial f(x) \cap N_A(x) \neq \emptyset$; ou equivalentemente, $0 \in \partial f(x) + N_A(x)$.

APÊNDICE D

NOÇÕES BÁSICAS DE TEORIA DA MEDIDA

Nesta seção, apresentamos algumas definições e resultados básicos da Teoria da Medida. As referências básicas podem ser encontradas em Folland [21], Fernandez [19] ou Bartle [2]. Seja X um conjunto nao vazio.

D.1 Sigma-Álgebra

Definição D.1. *(**Sigma-Álgebra**). Uma família $\mathcal{S}$ de subconjuntos de X é uma Sigma-álgebra se*

(i) $\emptyset, X \in \mathcal{S}$;

(ii) se $A \in \mathcal{S}$, então o complemento de A, denotado por A^c, pertence a $\mathcal{S}$;

(iii) se $\{A_n\}_{n=1}^{\infty}$ é uma sequência de conjuntos em $\mathcal{S}$, então a união $\bigcup_{n=1}^{\infty} A_n$ pertence a $\mathcal{S}$.

Definição D.2. *(**Sigma-álgebra de Borel-$\mathcal{B}$**). A Sigma-álgebra de Borel é sigma-álgebra gerada por subconjuntos abertos de $\mathbb{R}^n$.*

Definição D.3. *(**Conjunto Mensurável**). O par $(X, \mathcal{S})$, sendo X um conjunto não-vazio e $\mathcal{S}$ uma sigma-álgebra de X, é denominado de espaço mensurável. Qualquer conjunto em $\mathcal{S}$ é chamado de conjunto $\mathcal{S}$-mensurável, mas quando a sigma-álgebra é fixa (como geralmente é o caso), o conjunto será chamado de mensurável.*

Observação D.1. *Seja $X = \mathbb{R}$. A Sigma-álgebra de Borel é a sigma-álgebra gerada por todos os intervalos abertos (a, b) em $\mathbb{R}$. Observe que a sigma-álgebra de Borel $\mathcal{B}$ é também a sigma-álgebra gerada por todos os intervalos fechados $[a, b]$ em $\mathbb{R}$. Qualquer conjunto em $\mathcal{B}$ é chamado um conjunto de Borel.*

D.2 Função Mensurável

Consideramos um espaço mensurável fixo $(X, \mathcal{S})$, logo temos a seguinte definição:

Definição D.4. *(**Função Mensurável**). Seja $f : X \longrightarrow \mathbb{R}$ uma função real em X. Dada uma sigma-álgebra $\mathcal{S}$, diremos que a função f é mensurável (com respeito a $\mathcal{S}$) se para cada número α, o conjunto*

$$\{x \in X : f(x) > \alpha\} \tag{D.1}$$

é mensurável.

São consequências da definição de função mensurável:

(i) Os conjuntos $\{x \in X : f(x) = \alpha\}$, $\{x \in X : f(x) \geq \alpha\}$, $\{x \in X : f(x) < \alpha\}$ e $\{x \in X : f(x) \leq \alpha\}$ são mensuráveis.

(ii) Se $X = \mathbb{R}$ e $\mathcal{S}$ é a sigma-álgebra de Borel $\mathcal{B}$, então qualquer função contínua $f : \mathbb{R} \longrightarrow \mathbb{R}$ é Borel mensurável.

(iii) Se $X = \mathbb{R}$ e $\mathcal{B} = \mathcal{S}$, então qualquer função monótona é Borel mensurável.

D.3 Medida

Introduzimos a noção de espaço mensurável $(X, \mathcal{S})$ consistindo de um conjunto X e uma sigma-álgebra $\mathcal{S}$ de subconjuntos de X. Vamos considerar certas funções definidas em $\mathcal{S}$ e tomando valores no conjunto do números reais ou reais estendidos. Estas funções generalizam a ideia de comprimento, área, massa, assim em diante.

Definição D.5. *Uma função $\mu : \mathcal{S} \longrightarrow [0, +\infty]$ é chamada de **medida** em $(X, \mathcal{S})$ se*

(i) $\mu \geq 0$ para todo $A \in \mathcal{S}$;

(ii) $\mu(\emptyset) = 0$;

(iii) se $\{A_n\}_{n=1}^{\infty}$ *é qualquer sequência disjunta de conjuntos em* $\mathcal{S}$*, então*

$$\mu\left(\bigcup_{n=1}^{\infty} A_n\right) = \sum_{n=1}^{\infty} \mu(A_n).$$

Vamos apresentar um exemplo considerado importante:

Exemplo D.1. *Se* $X = \mathbb{R}$ *e* $\mathcal{B} = \mathcal{S}$*, a sigma-álgebra de Borel, então podemos mostrar que existe uma única medida* $\lambda : \mathcal{B} \to [0, +\infty]$ *a qual coincide com o comprimento dos intervalos abertos, isto significa que se* $A = (a, b)$*, então* $\lambda(a, b) = b - a$*. Esta única medida é geralmente chamada* ***medida de Lesbegue (ou Borel)****. Esta não é uma medida finita, mas sim* σ*-finita.*

Podemos apresentar alguns resultados simples da teoria da medida:

Lema D.1. *Seja* μ *uma medida definida em uma sigma-álgebra* $\mathcal{S}$*. Se* $E, F \in \mathcal{S}$ *e* $E \subset F$*, então* $\mu(E) \leq \mu(F)$*. Se* $\mu < +\infty$*, então* $\mu(F \setminus E) = \mu(F) - \mu(E)$*.*

Também temos:

Lema D.2. *Seja* $\mu : \mathcal{S} \to [0, +\infty]$ *uma medida. Temos as seguintes afirmações:*

a) Se $\{A_n\}$ *é uma sequência crescente em* $\mathcal{S}$*, então:*

$$\mu\left(\bigcup_{n=1}^{\infty} A_n\right) = \lim_{n\to\infty} \mu(A_n).$$

b) Se $\{B_n\}$ *é uma sequência decrescente em* $\mathcal{S}$ *e* $\mu(B_1) < +\infty$*, então*

$$\mu\left(\bigcap_{n=1}^{\infty} B_n\right) = \lim_{n\to\infty} \mu(B_n).$$

D.4 Espaço de Medida

Definição D.6. *(Espaço de Medida). É uma tripla* $(X, \mathcal{S}, \mu)$ *que consiste de um conjunto* X*, uma sigma-álgebra* $\mathcal{S}$ *de subconjuntos de* X *e uma medida* μ *definida em* $\mathcal{S}$*.*

Definição D.7. *Se* μ *é uma medida definida em* $\mathcal{S}$*, então uma relação envolvendo os elementos de* X *é dita valer* ***quase sempre*** *se o conjunto* A *de todos pontos para os quais a relação falha é um conjunto de medida nula, isto é,* $\mu(A) = 0$*.*

D.5 Integral de Lebesgue

Seja $\mathcal{S}$ uma sigma-álgebra em um conjunto X e μ uma medida positiva em $\mathcal{S}$. Se $s : X \to \mathbb{R}^+$ é uma função simples mensurável, da forma

$$s = \sum_{i=1}^{n} \alpha_i \mathcal{X}_{A_i},$$

onde $\alpha_1, \ldots, \alpha_n$ são valores distintos de s e $\mathcal{X}_{A_i}$ é a função característica do conjunto A_i. Se $E \in \mathcal{S}$ definimos a integral de s em E com relação a μ por:

$$\int_E s d\mu = \sum_{i=1}^{n} \alpha_i \mu(E \cap A_i).$$

Definição D.8. *Seja $f : X \longrightarrow \mathbb{R}^+$ uma função mensurável e $E \in S$. Definimos*

$$\int_E f d(\mu) = \sup_{s \leq f} \int_E s d\mu. \tag{D.2}$$

A integral do lado direito de (D.2) é chamada integral de Lebesgue de f sobre E, com relação à medida μ. Quando $\int_E f d(\mu) < \infty$, dizemos que f é integrável.
A integral para qualquer função mensurável é como segue:
$f : X \longrightarrow \mathbb{R}$ é integrável se,

$$\int_E f^+ d(\mu) < \infty \text{ e } \int_E f^- d(\mu) < \infty,$$

e

$$\int_E f d(\mu) := \int_E f^+ d(\mu) - \int_E f^- d(\mu)$$

onde $f := f^- f^-$.

Da definição de integral de Lebesgue temos as seguintes propriedades:

Teorema D.1. *(**Teorema da Convergência Monótona**). Seja $\{f_n\}$ uma sequência de funções mensuráveis em X, e suponha que*

a) $0 \leq f_1(x) \leq f_2(x) \leq \cdots < \infty$ para todo $x \in X$;

b) $f_n(x) \to f(x)$ quando $n \to \infty$, para cada $x \in X$. Então, f é mensurável e

$$\int_X f_n d(\mu) \to \int_X f d(\mu).$$

Lema D.3. *(**Lema de Fatou**). Seja $\{f_n\}$ uma sequência de funções não-negativas e e* *integráveis em* $(X, \mathcal{S}, \mu)$. *Então*

$$f(x) := \lim_{n\to\infty} \inf f_n(x)$$

é mensurável e

$$\int_X \left(\lim_{n\to\infty} \inf f_n\right) d(\mu) \leq \lim_{n\to\infty} \int_X f_n d(\mu).$$

Teorema D.2. *(**Teorema da Convergência Dominante de Lebesgue**). Suponha uma sequência de funções integráveis* $\{f_n\}$ *que converge para uma função* f *de valores reais mensurável, quase sempre. Se existir uma função integrável* g *tal que* $|f_n| \leq g$ *para todo* n, *então* f *é integrável e*

$$\lim_{n\to\infty} \int_X f_n(x)d(\mu) = \int_X f(x)d(\mu).$$

Teorema D.3. *(**Teorema de Fubini**). Sejam* $(X_1, \mathcal{S}_1, \mu_1)$ *e* $(X_2, \mathcal{S}_2, \mu_2)$ *dois espaços mensuráveis sigma-finitos. Se* f *é uma função mensurável* $\mathcal{S}_1 \times \mathcal{S}_2$ *em* $X_1 \times X_2$ *e não-negativa, então a seguintes integrais são iguais*

$$\begin{aligned} \int_{X_1\times X_2} f(x,y)d(\mu_1 \times \mu_2) &= \int_{X_1} \left(\int_{X_2} f(x,y)d(\mu_2)\right) d(\mu_1) \\ &= \int_{X_2} \left(\int_{X_1} f(x,y)d(\mu_1)\right) d(\mu_2). \end{aligned}$$

D.6 Outros resultados

Definição D.9. *Uma função* $f : [a, b] \to \mathbb{R}$ *é absolutamente contínua se, satisfaz-se uma das seguintes:*

1. $\forall \epsilon > 0 \, \exists \, \delta > 0$ *tal que para cada coleção* $\{[a_i, b_i]\}$ *de sub-intervalos disjuntos de* $[a, b]$ *tem-se*

$$\sum_i (b_i - a_i) < \delta \Rightarrow \sum_i |f(b_i) - f(a_i)| < \epsilon.$$

2. $\exists \, v \in L(a, b)$ *(o espaço das funções integráveis definidas em* $[a, b]$*) tal que*

$$f(t) = f(a) + \int_a^b v(\tau)d\tau, t \in [a, b].$$

Teorema D.4. *(Teorema fundamental do cálculo para integrais de Lebesgue). Se* $-\infty < a < b < \infty$ *e* $F : [a, b] \to \mathbb{R}$, *são equivalentes:*

a) F *é absolutamente contínua em* $[a, b]$.

b) $F(x) - F(a) = \int_0^x f(y)dy$ *para algum* $f \in L_1$.

c) F *é diferenciável q.t.p. em* $[a, b]$, $F'(x) \in L_1$ *e*

$$F(x) - F(a) = \int_a^t F'(y)dy.$$

Teorema D.5. *Seja* G *uma função contínua e crescente sobre* $[a, b]$ *e* $G(a) = c$, $G(b) = d$. *Se* f *é uma função Borel mensurável e integrável sobre* $[c, d]$, *então*

$$\int_c^d f(t)dy = \int_a^b f(G(x))dG(x),$$

e

$$\int_c^d f(t)dy = \int_a^b f(G(x))G'(x)dx$$

se G *é absolutamente contínua.*

Teorema D.6. *(Teorema de Lusin). Seja* $f : [a, b] \to \mathbb{R}$ *uma função mensurável à medida de Lebesgue. Então para todo* $\delta > 0$, *existe uma função* $g : [a, b] \to \mathbb{R}$ *contínua, tal que:*

$$\mu\left(\{x \in [a, b] : f(x) \neq g(x)\}\right) < \delta.$$

Autorizo a reprodução xerográfica para fins de pesquisa.

São José do Rio Preto, 03 de Março de 2016.

Raul Tintaya Marcavillaca